Manipulating DNA repair for improved genetic engineering in Aspergillus

Nødvig, Christina Spuur

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Manipulating DNA repair for improved genetic engineering in *Aspergillus*

Christina Spuur Nødvig

PhD Thesis

Technical University of Denmark

Department of Systems Biology

Preface

This thesis is the culmination of 3 years and 4 months of my PhD studies, taking place from July 2012 to November 2015 in the Eukaryotic Molecular Cell Biology group, Section of Eukaryotic Biotechnology, at the Department of Systems Biology, at the Technical University of Denmark. Financial support was provided by the MycoFuelChem project, grant 11–116803 from The Danish Council for Strategic Research.

This have be an exciting time with ups and downs, and several persons have really helped me along this journey and deserves thanks. First I would like to thank my supervisor Uffe Hasbro Mortensen for all the scientific discussion, for giving me the academic freedom to grow as a scientist, and for being there for me when I needed support or guidance. Also Jakob Blæsbjerg Hoof deserves big thanks. While not my supervisor on the official paper, he has nonetheless been an excellent supervisor, teaching me all the *Aspergillus* tricks in the lab and been there for me, both for valuable discussion and moral support.

Many thanks to Maria Lund Nielsen, for being the best office mate, it has been great fun and lots of both great or crazy ideas have come out from our office. Along with Maria and Jakob, I would also like to thank Diana Chinyere Anyaogu, Zofia Dorota Jarczynska, Paiman Khorsand-Jamal and Martin Schalén for all the good times in lab. Special thanks to Gustav Hammerich Hansen for being such a good friend, always ready to cheer me up when I needed it the most or to celebrate even the little victories with. In general all the wonderful people in the Eukarotic Molecular Cell Biology group and rest of the Section of Eukaryotic Biotechnology deserves a big thanks for providing an excellent and inspiring work environment with room for fun as well.

Finally I would like to thank my father in law, Steen Bennike Mortensen for careful and critical reading of my thesis and proving valuable feedback. A big thank to the rest of my family as well for all their love, support and understanding throughout my studies. Last but definitely not least, I would like to thank my beloved husband Benjamin Bennike Aagren, for putting up with me through three stressful years, his continuous love and support have been beyond measure.

2nd of November 2015

Christina Spuur Nødvig

Summery

Aspergillus is a genus of filamentous fungi, which members includes industrial producers of enzymes, organic acids and secondary metabolites, important pathogens and a model organism. As such no matter the specific area of interest there are many reasons to perform genetic engineering, whether it is metabolic engineering to create better performing cell factory, elucidating pathways to study secondary metabolism etc. In this thesis, the main focus is on different ways to manipulate DNA repair for optimizing gene targeting, ultimately improving the methods available for faster and better genetic engineering strategies.

Chapter 1 gives an introduction to the genus Aspergillus and some of the tools relevant to fungal genetic engineering. It also contains a short introduction to DNA repair and its interplay with gene targeting and finally an overview over the different genome editing technologies, providing a background for the other chapters.

Aspergillus nidulans is a model organism, with a range of genetic tools developed, and therefore the approach in this thesis has been to use it for proof of concept, and once a method has been established in *A. nidulans* it can be transferred to other Aspergilli.

In **chapter 2**, the focus is on a concept for allowing simultaneous, but transient disruptions of genes of interest, with the main goal being creating a strain with transient disruptions of both *pyrG* and *nkuA*. This would yield a strain with a robust selection marker and high rates of gene targeting. However, since the same trait, while beneficial for genetic engineering, is a detriment during a fermentation process where a robust DNA repair system and a prototrophic background are important traits. By inserting a marker gene into an intron, flanked by loxP sites, its presence will interfere with intron splicing, however it can be excised by the cre recombinase, leaving only a single loxP scar in the intron, which if placed correctly will not interfere with intron-splicing and restore function. This was demonstrated to work in two different introns in the pigment gene *yA* in *A. nidulans* as a proof of concept, and the concept was expanded to include the *tetON* promoter controlling the cre recombinase gene as part of the insert and also testing whether it is tolerated to insert or replace an intron from one gene to another, which were partial successful. Finally, a transient *pyrG* mutation was introduced in a marker free strain of *A. nidulans,* and successfully tested and similarly a disruption of *nkuA* was made.

Genome editing is not a new concept, but never has it been as accessible as it is now due to the CRISPR-Cas9 technology. In **chapter 3**, a versatile CRISPR-Cas9 system for use in various Aspergillus species was made, consisting of four vectors each with a different selection marker. To successfully express the two components of CRISPR-Cas9, Cas9 was codon-optimized to *A. niger*, and a ribozyme based strategy was used for gRNA expression. With a functional system in place, I demonstrated how it could be used to disable genes by mutagenesis, but also how it could be used to greatly enhance gene targeting frequencies in wild type strains, similar to what is observed when using non-homologous end-joining (NHEJ) deficient strains. Furthermore, a perl script for identifying protospacers in common for gene homologs across multiple species was developed. While the initial experiments were made in *A. nidulan*s and *A. aculeatus*, this was used to quickly demonstrate that the system could be used in more species, and mutagenesis was done in four additional species.

In **chapter 4** the focus stays on CRISPR-Cas9, focusing on three aspects. Alternative methods for gRNA expression, strategies for limiting off-targeting effects, and how to combine CRISPR-Cas9 with the traditional strategy of disabling NHEJ for even greater results. Four other promoters were tested for their ability to express functional gRNA, chosen based on which has been shown to work in other species. However, none of them worked in *A. nidulans* and the ribozyme based strategy remained the most effective. Next two strategies for limiting potential off-targeting effects using CRISPR-Cas9 were explored. One was based on using a shorter protospacer than the 20 bp that is normally used. 17 bases of length was tested in *A. nidulans*, but lead to aberrant sporeless colonies. The other strategy was based on inactivating one of the cleavage domains in the Cas9 protein, turning it into a nickase and then use two gRNAs to create two nicks in close proximity on opposite strands rather than a single double-strand break. Results indicated that it could be a viable strategy to stimulate gene targeting while lowering potential off-targeting. Finally I demonstrated by combining CRISPR-Cas9 with an NHEJ deficient background, it was possible to do gene targeting without selection for DNA to be integrated and furthermore that, for small changes, such as for introducing point mutations, it was possible to use short single-stranded oligos.

Overall several very useful tools for genetic engineering of various Aspergillus species were developed as a part of this thesis, and especially the CRISPR-Cas9 based tools have the potential to transform genetic engineering strategies.

Sammenfatning

Aspergillus er en slægt filamentøse svampe, som inkluderer flere industrielt relevante arter, som er i stand til at producere enzymer, organiske syrer og sekundære metabolitter, men også arter som er vigtige patogener eller model organismer. Derfor, uanset det specifikke interesseområde, er der mange årsager til, hvorfor det kan være relevant at genmanipulere disse organismer, hvilket inkluderer at lave cellefabriker, undersøge specifikke synteseveje, studere sekundær metabolisme el.lign. Denne afhandling er centreret omkring forskellige måder, DNA reparations mekanismerne kan blive udnyttet til at optimere introduktion af specifikke mutationer og generelt forbedre strategier for at genmanipulere Aspergillus arter.

Kapitel 1 er en introduktion til Aspergillus slægten og mange af de relevante redskaber der er til genmanipulation af disse skimmelsvampe. Derudover er der også en kortfattet introduktion til DNA reparation og dets relevans for genmanipulation. Endelig er der en kort oversigt over nogle af de forskellige gen editerings teknologier som eksisterer, hvilket er relevant for de senere kapitler.

Aspergillus nidulans er model organisme, for hvilken adskillige genetiske redskaber er blevet udviklet. Derfor har udgangspunktet i denne afhandling været at *A. nidulans* har været anvendt til at teste koncepter, hvorefter de bliver overført til andre Aspergillus arter.

I **kapitel 2**, var målet en metode for at lave reversible gen afbrydelser, hvoraf hovedformålet var at lave en stamme midlertidigt afbrudt i både *pyrG* og *nkuA* i *A. nidulans*. Dette vil resultere i en stamme som har en mod-selekterbar genetisk selektions markør, men som også inkorporerer heterolog DNA ved hjælp homolog rekombination, hvilke tillader at lave præcise ændringer i svampens genom. De samme to træk er dog en ulempe i fermenterings processer, hvor en god produktionsstamme gerne skal være prototrof og med et robust DNA reparations system. Ideen er at indsætte en dominant markør flankeret af loxP sites ind i intronen af det gen som skal afbryde. Så længe det er indsat i intronen vil dets tilstedeværelse forhindre intron splejsning, i praksis forhindre translation. Til gengæld med cre rekombinasen er det muligt at skære det indsatte DNA ud igen og kun efterlade et enkelt loxP site, hvilket, hvis det er placeret korrekt, ikke interfererer med intron splejsning. Først blev konceptet testet ved succesfuldt at afbryde og restaurere pigment genet *yA*. Derefter blev konceptet udvidet til at inkludere cre rekombination genet under kontrol af tetON promoter, og derudover blev det testet om det muligt at indsætte eller udskifte en intron fra et gen til

et andet, hvilket var delvist succesfuldt. Til sidst blev *pyrG* succesfuldt midlertidigt afbrudt i en markørfri stamme, og en tilsvarende mutation blev indført i *nkuA* genet.

Genom editerings teknologi er ikke et nyt koncept, men det har aldrig været så tilgængeligt som det er nu. I **kapitel 3** blev et fleksibelt CRISPR-Cas9 system udviklet, bestående af fire vektorer med forskellige selektions markører, hvilket kan blive anvendt i forskellige arter. CRISPR-Cas9 kræver to komponenter for at virke. Cas9 var kodon optimeret til at matche *A. niger*, mens en strategi baseret på ribozymes blev anvendt til at udtrykke gRNA'et. Med et funktionelt system, var det muligt at afbryde gener ved mutagenese, men også booste frekvensen hvorved heterolog DNA bliver integreret ved homolog rekombination, tilsvarende den effekt som ses når man fjerner svampens evne til at reparere DNA dobbelt strengs brud ved hjælp af illegitim rekombination. Derudover blev et Perl script udviklet til at kunne identificere protospacers som fungerer i flere homologer af det samme gen i forskellige arter. Som udgangspunkt blev CRISPR-Cas9 testet i to arter, *A. nidulans* and *A. aculeatus*, men scriptet blev udnyttet til at demonstrere at systemet var funktionelt i yderligere fire arter.

Kapitel 4 fokuserer på tre forskellige aspekter af anvendelsen af CRISPR-Cas9. Først blev alternative promotere testet for deres evne til succesfuldt at udtrykke gRNA'er. Disse var valgt, baseret på hvad der har virket i andre organismer. Desværre virkede ingen af dem ud over den ribozyme baserede strategi præsenteret i kapitel 3. Derefter blev to forskellige strategier som kan begrænse potentielle problemer med "off-targeting" testet. Den første var baseret på at anvende protospacere som var 17 baser lange i stedet for 20, men det gav sære sporeløse kolonier og var ikke en succes. Derimod virkede strategien med at omdanne Cas9 til en nickase og anvende to gRNA'er til at lave to nicks i stedet for et enkelt dobbelt strengs brud, til at stimulere integration ved hjælp af homolog rekombination. Endelig viste jeg at man ved at kombinere CRISPR-Cas9 med at afbryde illegitim rekombination, er det muligt at lave gen modificationer uden at efterlade en selektion markør i genomet. Derudover demonstrerede jeg at det er muligt at bruge enkelt strengede oligoer i stedet for konventionelle vektorer når der skal introduceres små modifikationer.

Alt i alt har denne afhandling resulteret i adskillige brugbare værktøjer til at lave genetiske modifikationer og især de CRISPR-Cas9 baserede redskaber har potentiale til at ændre konventionelle strategier for genmanipulation i *Aspergillus* arter.

Publications

Peer-reviewed papers

- Nødvig CS, Nielsen JB, Kogle ME, Mortensen UH (2015) A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi. PLoS ONE 10(7): e0133085.

Conference contribution

- Nødvig CS, Nielsen JB, Kogle ME, Mortensen UH. A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi. Poster and oral presentation at 28th Fungal Genetics Conference, Asilomar, CA, USA (2015)

- Nødvig CS, Buron LD, Strucko T, Jarczynska ZD, Mølgaard L and Mortensen UH. Controlled gene amplification enables high, stable, selection free gene expression in *S. cerevisiae.* Poster presentation at Synthetic Bioology, Engineering, Evolution and Design Conference, Los Angeles, CA, USA (2014)

- Nødvig CS, Buron LD, Strucko T, Jarczynska ZD, Mølgaard L and Mortensen UH. Controlled gene amplification enables high, stable, selection free gene expression in *S. cerevisiae.* Poster presentation at 8th Danish Conference on Biotechnology and Molecular Biology (2013)

- Nødvig CS, Buron LD, Strucko T, Jarczynska ZD, Mølgaard L and Mortensen UH. Controlled gene amplification enables high, stable, selection free gene expression in *S. cerevisiae.* Poster presentation at 4th International Conference for Biomolecular Engineering, Fort Lauderdale, FL, USA (2013).

List of Abbreviations

AspGD	Aspergilluls genome Database
BIR	break induced replication
bp	basepair
CRISPR	Clustered regular interspaced palindromic repeats
crRNA	CRISPR RNA
dHJ	double Holliday junction
DSB	Double strand break
DSBR	Double strand break repair
gDNA	genomic DNA
gRNA	guide RNA
HR	Homologous recombination
kDa	kilo Dalton
MMEJ	Microhomology mediated end-joining
NHEJ	Non-homologous end-joining
NLS	nuclear localization signal
nt	nucleotides
RPA	Replication protein A
SDSA	synthesis-dependent strand annealing
sgRNA	single guide RNA
ss	single stranded
SSA	single strand annealing
TALE	Transcription activator-like effector
TALEN	Transcription activator-like effector nuclease
tracrRNA	Trans-activating CRISPR RNA
ZF	Zinc-finger
ZFN	Zinc-finger nuclease

Table of contents

Chapter 1: Introduction

Industrial fungal biotechnology

The fungal kingdom comprises of a wide range of very diverse organism. Members includes yeasts, such as *Saccharomyces cerevisiae*, mushrooms and filamentous fungi. In nature fungi are important degraders of biomass [1], able to grow in a wide range of environments.

Aspergillus is a genus of filamentous fungal species, containing a lot of scientifically, medically or industrial important species [2]. *Aspergillus nidulans* was early proposed as a model organism for genetics [3] and later also used for studies of the cell cycle and mitosis in eukaryotes [4]. *Aspergillus oryzae* has a much longer history of human use, being used for a wide range of Asian fermented products such as soy sauce, miso and sake for more than 2000 year [5], but is also used in modern biotechnology for the production of enzymes, and is considered an industrially relevant fungus. *Aspergillus niger* is likewise used in modern biotech, both for the production of enzymes, but also for production of organic acids, such as citric acid. Together with *Aspergillus fumigatus*, an important fungal pathogen, infecting immunocompromised patients [6], these four aspergilli were the first to be genome sequenced, though the *A. niger* genome was released later [7–10], and although many others have joined since then, and currently approximately 300 Aspergilli are being or stated to be sequenced, as a part of the Aspergillus Whole-genus sequencing project (http://genome.jgi.doe.gov/). While *A. niger* and *A. oryzae* are the two big industrial Aspergilli, and both of them have obtained GRAS status from the FDA, many of the traits that makes them excellent production organisms, can also be found in other Aspergilli. Due to the saprophytic lifestyle, Aspergilli in general have a strong secretion apparatus, and are able to secrete the enzymes necessary for degrading of biomass. Besides the potential of secretion of proteins and organic acids, Aspergilli also presents a huge reservoir of interesting secondary metabolites. For instance *Aspergillus terreus* is a natural producer of lovastatin, an important cholesterol lowering agent. The secondary metabolites of some species also present huge problems. For instance *Aspergillus flavus* being a food spoilage agent, but also producer of the extremely carcinogenic compound aflatoxin. There are therefore many compelling reasons for studying the various Aspergilli in detail, also at a molecular level, both for basic research purposes and for the improvement of industrial strains. With more and more sequences becoming available for study, the pressure is on the molecular tools for genetic engineering to keep up.

Tools for genetic engineering of filamentous fungi

The first report on a transformation in a filamentous fungi, was on *Neuraspora crassa* [11], while a report on using *Aspergillus nidulans* followed in 1983 [12]. The fungal cell wall provided a challenge, unlike when transforming prokaryotes. This was however overcome by protoplastation, in which the cell wall is enzymatically degraded. Other methods for transformation includes electroporation [13], biolistic transformation [14] and Agrobacterium-mediated transformation [15]. Once the DNA is introduced, there are essentially two ways for it to be propagated, either by integration into the genome or by autonomously replication, such as by the AMA1 element [16]. While the AMA1 element can be used for transient expression, it is mitotically unstable, being readily lost over time [16], and in order to be used for stable expression, genomic integration is necessary. The manner in which DNA can integrate into the genome depends on several factors and will be discussed in the next section.

Another factor that is crucial for transformation is the availability of selection markers, in order to separate cells which have taken up the DNA in question from those who have not. Selection markers can generally be divided into nutritional and antibiotic markers. Various antifungal drugs have successfully been used as selection markers in various Aspergilli, including hygromycin [17,18], bleomycin [19], and oligomycin [20]. Nutritional markers include auxotrophic markers which are created by disabling a gene in a pathway responsible for the synthesis of an essential compound. Commonly used markers includes *pyrG*, encoding an orotidine-5´-phosphate decarboxylase, which if disabled causes uridine auxotrophy [21] and *argB* encoding an ornithine carbamoyltransferase which if disabled causes arginine auxotrophy, but many other nutritionally important genes can similarly be used. Many of these being responsible for the synthesis of various amino acids or vitamins. Of special interest is that while some nutritional genes behave as a classical auxotrophic marker gene in their native host, requiring a null-mutant before it can be utilized as selection marker, others can also be used as a dominant nutritional selection marker in other species. One example of this type of marker gene is the *A. nidulans amdS* gene, which if disrupted or deleted prevents the utilization of acetamide as nitrogen source [22]. Wildtype *A. niger* however grows poorly with acetamide as nitrogen source, and as such *amdS* can be used a dominant nutritional selection marker [23]. In addition some selection markers, including *pyrG* and *amdS* have the advantage that not only is it possible to select for their presence but also their absence. The *pyrG* marker can be counter-selected by 5-fluoroorotic acid (5-FOA), while *amdS* can be counter-

selected by 5'-fluoroacetamide (FAA). This allows these marker genes, if flanked with repetitive DNA sequences, to be reused since the marker genes then can be recycled by direct repeat recombination.

Another important consideration, especially for expression of heterologous proteins, but also as a tool for changing regulation of native genes, is the choice of promoter since it affects both the regulation and expression levels of the gene it controls. Promoters are in general divided into constitutive and inducible promoters. While constitutively expressed genes are active independent on growth conditions, genes expressed from inducible promoters can potentially allow for fine-tuning of the fermentation process by controlling when the protein or metabolite production starts.

Some of the first inducible promoters used includes the *A. nidulans alcA* promoter [24] from the alcohol dehydrogenase I, and the glucoamylase, *glaA* of *A. niger*. The alcA promoter is induced by the presence of acetaldehyde, while the *glaA* promoter is induced when maltose or starch is the sole carbon source [25]. A drawback is that both of these promoters are repressed by glucose which limits their application. In general the considerations for choosing an inducible promoter includes both the strength of the promoter when induced, the tightness of the promoter when not induced, the suitability of inducers and repressor to the fermentation process and the potential cost of inducers. Besides *alcA* and *glaA* there are many other Aspergillus promoters, responding to different environmental condition or inducers, which have successfully been used for various production strategies [2,25]. An alternative to using native Aspergillus promoters, is adapting inducible promoter systems from other organisms, which often depends on different types of inducers, and is de-coupled from the metabolism of the cell, such as the prokaryotic *tetON* promoter, responding to tetracycline and doxycycline [26,27].

Alternative to using inducible promoters, is constitutive promoters which are not dependent inducers or tied to the growth condition. Classical examples of those includes the *A. nidulans gpdA* promoter, from the Glyceraldehyde-3-phosphate dehydrogenase gene [28], the *A. oryzae tef1* promoter, from the translation-elongation factor gene [29] and the *A. nidulans oliC* promoter, from a mitochondrial ATP synthase subunit gene. Interestingly most of these promoters originating from one Aspergilli, functions across many of the different Aspergilli species.

While also impacting transcription, often much less consideration is given to the choice of terminator [2]. However the terminator affects the stability of the transcript and thus the half-life of

the RNA [30], and as such it can have significant effect on protein production. For instance in the yeast, *S. cerevisiae* exchanging the commonly used *CYC1* terminator for a synthetic one, increased production of yellow fluorescent protein by 3.7 fold, while the promoter was the same [31].

Another very important factor greatly affecting gene expression is where and in how many copies a gene is located. When DNA is integrated into the genome it can happen by three possible DNA repair pathways, as described in the next session, however it comes down to where the DNA goes in at a defined location by homologous recombination (HR) or to a random or potential several random locations by non-homologous end-joining (NHEJ). Due to its relatively late recognition as an independent DNA repair pathway, to which degree events classified NHEJ events are actually a results of micro-homology end-joining (MMEJ) is unknown. The location where the DNA is integrated has a huge impact on gene expression. Some locations support high levels of gene expression, while for instance chromatin structure, in the form of hetero chromatin and telomeric silencing effects can negatively impact transcription [32]. Another aspect to consider besides the impact of the location of the gene in question, is the effect on the neighboring genes. A common strategy to circumvent these issues is by the definition of integration sites, where typically identical constructs for expression of an easily quantifiable protein is tested in different location, combined with test of either growth rate or expression of neighbor genes [33–35]. Of course to take advantage of such predefined sites, it requires that the location of the insert can be controlled, which requires integration by HR. Also other sorts of modifications, focusing on the existing genes, such as gene deletions and point mutations are fully dependent on integration by HR. However, from the perspective of insertion of genes, random integration is still interesting since it comes with both advantages and disadvantages. Random integration of DNA poses the risk of disrupting native genes or hitting locations detrimental for gene expression, but there is also the possibility of multiple inserts. Data from *A. oryzae*, *A. niger* and *Trichoderma reesei* suggest that up to approximately five copies there is linear correlation between copy numbers of an insert and resulting protein production, while dose/response relation is lacking at higher copy numbers [36]. One reason for the limited effects of multiple copies, in the case of inducible promoters such as alcA and glaA promoters are limited availability of the regulation proteins [2,25,36]. Another concern is a tendency of multiple inserts to integrate in tandem, resulting in genomic instability due to the risk of direct repeat recombination.

DNA repair

Genetic engineering is a wide concept, which roughly can be defined as techniques that allow for editing of the genome of an organism and includes introduction of new genes, either targeted or randomly inserted, deletion or disruption of existing genes and the introduction or repair of point mutations.

The ability to do targeted genetic manipulations is intrinsically linked to DNA repair in the cell. Eukaryotes have traditionally been considered to have two major repair pathways, non-homologous end-joining (NHEJ) and homologous recombination (HR), while more recent microhomology-mediated end-joining (MMEJ) have begun to be considered as a third major repair pathway [37], rather than a backup system to NHEJ. The choice between these three pathways has major impact on how amenable an organism is to genetic engineering, and depends both on the organism itself, but also in which stage of the cell cycle it is in. The three pathways are illustrated in Figure 1.1.

Relatively little research have been done on DNA repair in *Aspergillus*, and therefore, unless specifically stated the following sections are based on information from the yeast *S. cerevisiae* or mammalian systems, since significantly more research into DNA repair are being done in these organisms. Furthermore, while there are many types of DNA damage and also many ways to repair it, this section will focus on double strand break repair, since those breaks are the interesting ones concerning gene targeting and genome editing technologies.

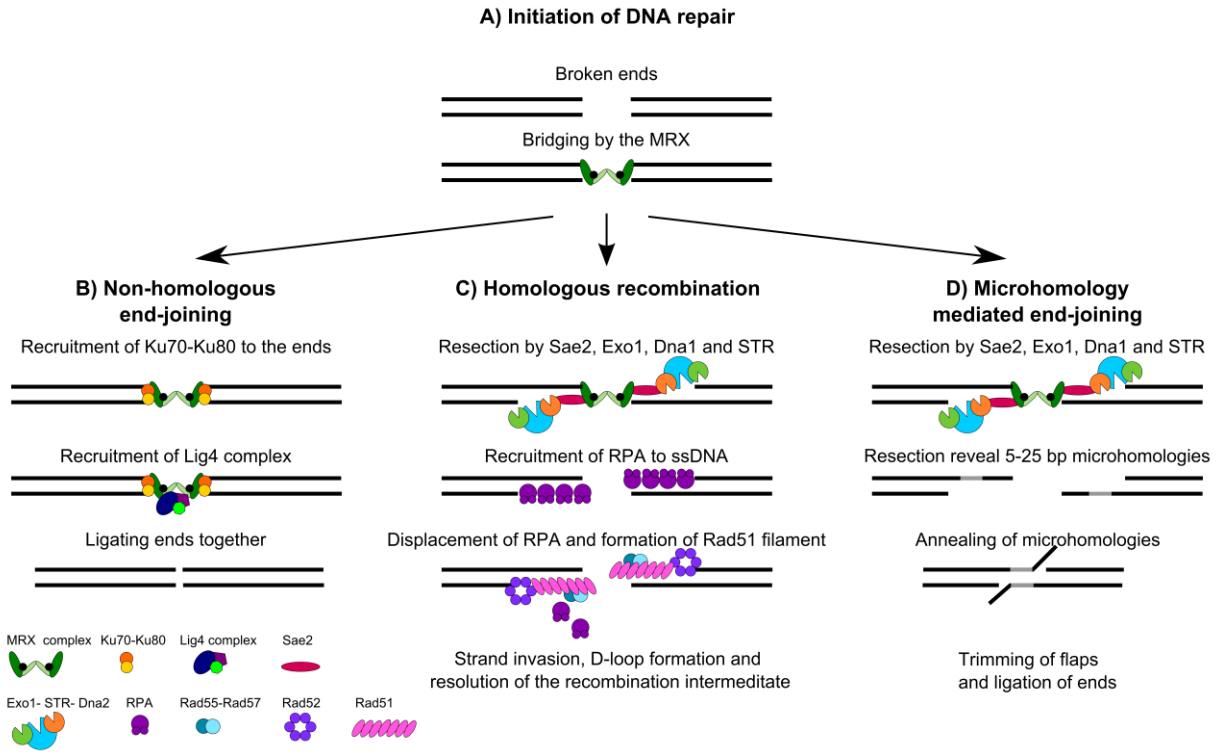

Figure 1.1 Overview of the three major DNA repair pathways. A) Independent of pathway choice, repair is initiated by the recruitment of the MRX complex, bridging the broken ends. B) NHEJ is initiated by the recruitment of the Ku70-Ku80 hetero dimer, protecting the broken ends. The Lig4 complex is then recruited to the site and the ends are ligated together. C) HR by contrast is initiated by resection, creating 3' overhangs. RPA is recruited to the ssDNA, and then it is displaced by Rad51, assisted by Rad52. Rad51 forms a filament assisted by Rad55 and Rad57, which can then perform strand invasion. D) MMEJ is similar to HR initiated by resection. The resection reveals microhomologies which can anneal to each other, resulting in a deletion the intervening sequence.

Non-homologous end-joining

NHEJ is initiated by the recruitment of the Ku70 and Ku80 proteins, which forms a heterodimer, which binds to the DNA ends, protecting them from further damage and assists in recruiting the proteins necessary for downstream activity [37]. The names refers to the size of the human variants of the proteins, with Ku70 being ≈ 70 k dalton (kDA) and likewise the Ku80 being ≈ 80 kDa. The NHEJ pathway was first discovered in eukaryotes, where the Ku proteins, while displaying significant differences in sequences, are functionally conserved, to the degree where the human or Drosophila Ku70 can complement Ku70 null mutants in yeast [38]. Many bacteria however, also possess an NHEJ repair pathway [39], but the bacterial Ku proteins generally have one domain less than their eukaryotic counterparts [40]. As mentioned above the Ku hetero dimer covers the ends, which helps recruitment of the next proteins in the NHEJ pathway, however the Ku proteins also have other roles besides DNA repair by NHEJ. In yeast they have been shown to be involved in telomere maintenance [41], which has shown to be the same also in many other eukaryotes [42].

Besides the Ku hetero dimer, two other protein complexes are required for repair by NHEJ. In yeast the MRX complex, consisting of Mre11, Rad50 and Xrs2, which unlike Ku also are involved in repair by HR [43]. Ku and MRX are recruited independently from each other, however it have been observed that in the absence of Ku MRX binds later [43]. MRX has several functions in DNA repair. It functions both as sensor of DSB, bridges the ends together, is involved in 5'-3' end resection and also plays a role in telomeric maintenance [37]. After MRX and the Ku hetero dimer have bound to the DNA strands the lig4 complex, consisting of lig4, lif1 and nej1 is recruited [37]. Interestingly while recruitment of the lig4 complex is dependent on the Ku hetero dimer, later dissociation is dependent on MRX [43]. The lig4 complex is responsible for ligating the ends together, often without any mutations introduced.

One important difference between NHEJ in yeast and in vertebrates, is that in vertebrates the Ku heterodimer is part of a DNA dependent protein kinase, with a catalytic subunit (DNA-PKcs), which is required for NEHJ [44]. Yeast however does not have a DNA-PKcs [44]. This means at least in *S. cerevisiae* that NHEJ is dependent on ends being compatible [37].

Homologous recombination

On a principal level, what distinguishes homologous recombination from the other two repair pathways is that instead of ligating available DNA ends together, it relies on a homologous donor sequence which is used a template for repair. Overall HR happens in three stages, resection, followed by strand invasion and then resolution of the recombination intermediate [45].

Limited resection is initiated by the MRX complex together with an endonuclease Sae2, which together can remove oligonucleotides from the 5' strand [37]. The MRX complex then assists recruiting of Exo1, Dna2 and a complex of Sgs1, Top3 and Rmi1, which all are involved in a more extensive resection [37].

Resection exposes long single-stranded (ss) DNA overhangs which is required for homology search and which is then covered by Replication Protein A (RPA) [45]. Next RPA interacts with Rad52 which are involved in displacing RPA and recruiting Rad51 to the site instead [46]. Rad51 forms a filament, which is stabilized by the Rad55-Rad57 dimer [47]. The Rad51 filament then invades a homologous donor, creating a displacement (D) loop [48], assisted by Rad54 [49], in which one of the invading strands displaces one of the donor strands and pairs with the other.

Depending on the initial conditions, this D-loop can then be resolved in different manners (Figure 1.2), but in general the mechanism is that the invaded strand is used as template to extend the invading strand by DNA synthesis and the process is finished by gap filling and ligation [46], however the involved protein machinery will not be described. The D-loop can be resolved in different ways, depending on several conditions. For a DSB with two ends there is in general two ways to resolved it. In synthesis-dependent strand-annealing (SDSA) only a single strand invades, and after extension it dissociates from the D-loop resulting in non-crossover event in which the donor sequence is unchanged [45,50]. In double-strand break repair (DSBR) the displaced strand is then invaded by the other overhang, resulting in a double holliday junction (dHJ). From that point one option is a dissolution of the dHJ which results in non-crossover event or the resolution of the dHJ which can result in either a crossover event or a non-crossover event [45]. A third mode of repair is break-induced replication (BIR) assisted by Rad54 which happens in the case of a DSB with only a single end, which typically happens as a result of a collapsed replication fork [45]. In BIR one strand invades and both the invaded and the displaced strand are used as template for repair in a manner similar to leading and lagging strand in DNA synthesis. This results in the entire chromosome arm being replicated, potentially leading to chromosomal rearrangement or loss of heterozygosity [45].

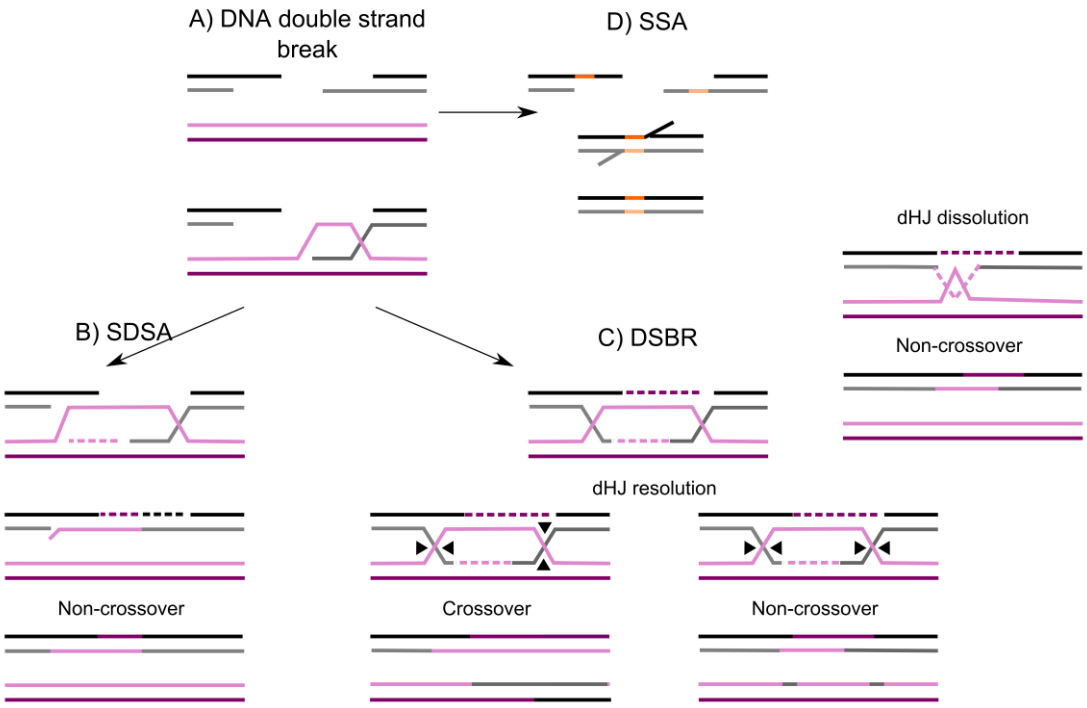

Figure 1.2 Models for DNA DSB repair by HR. A) After resection one strand invades and forms a D-loop. B) Synthesis-dependent strand annealing. A single strand invades and after extension, dissociates from the D-loop causing a non-crossover event. C) Double strand break repair. One strand invades, creating the D-loop, and then the displaced strand is also invaded forming a dHJ. The dHJ can be resolved in two ways, causing either a cross-over or a non-crossover event depending on the resolution. Alternatively a dissolution of the dHJ can happen, resulting in a non-crossover event. D) Single strand annealing. No D-loop formation happens. Instead resection reveals repeats which invades each other, resulting in the loss of the intervening DNA.

Two other scenarios are of particular interest. Single strand annealing (SSA) is the model that accounts for recombination between two direct repeats and which is the mechanism that allows excision of counter-selectable markers flanked by direct repeats. SSA unlike the other models for HR does not involve the formation of the D-loop, but instead each one of the repeats that are exposed by resection can pair with the other leading to the loss of one repeat and the intervening sequence [50].

The other scenario of interest is that of what happens during gene targeting, specifically what is called ends-out recombination, in which a double ended substrate with flanks of homology ends up substituting what is between the sequences in the chromosomal location matching the flanks. Both ends invade the homologous regions [51], however the exact method of resolution is not clear and might also be species dependent.

Microhomology mediated end-joining

Microhomology mediated end-joining, is also known as alternative non-homologous end-joining, since it was originally thought to be a backup mechanism for NHEJ, when classical NHEJ was unavailable. Unlike NHEJ, MMEJ is neither dependent on Ku or Lig4 [52]. MMEJ is initiated by resection similar to HR, by many of the same proteins as HR [52], and like both of the other two pathways MRX is involved. Where the two pathways separate is when it comes to RPA. Resection can reveal microhomologies of 5-25 bps [45] which can spontaneously anneal, however this spontaneous annealing is prevented by RPA [52]. Mechanically MMEJ a bit similar to SSA. MMEJ is a much more mutagenic pathway than HR or NHEJ, since it always results in deletion of the bp between the microhomologies, and it can even lead to translocation [37].

Choice of repair pathway

Several factors affect the choice of repair pathway in the case of a DSB. One of the big ones is the species in question, and also the cell type in multicellular species [53]. Another important factor is the cell cycle. In G1 phase resection seems to be reduced [37], favoring NHEJ, while resection is active in S and G2 phase which corresponds to those phases in which sister chromatid is present [45]. Another major factor is the type of break to be repaired, for instance modified ends can present a hindrance for the Ku heterodimer and such repair by NHEJ [37].

Manipulating DNA for improved gene targeting

The yeast *S. cerevisiae* is a prime example of an organism very amenable to genetic engineering, since when introducing extraneous DNA it will be integrated by HR with very high frequencies, even with very short flanks of homology [54]. Unlike *S. cerevisiae*, the various Aspergilli is not nearly is as amenable, integrating foreign DNA primarily by NHEJ, even if using long (1000-2000 bp) flanks of homology. With wildtype gene targeting frequencies ranging from 0% to 40%, it was a huge game changer for the ability to do genetic engineering in many species of filamentous fungi, including Aspergilli, when it was discovered that the disabling of one of the unique components in the NHEJ pathway, could greatly enhance gene targeting frequencies [55–58]. Viable targets for

disruption includes Ku70, Ku80 or LigD, and when one of these are disabled, greatly enhanced gene targeting frequencies are observed, In many cases close to a hundred percent depending on flank length of the gene targeting substrate and the specie in question. Similarly it has been attempted to increase gene targeting frequencies by upregulating the HR pathway rather than disabling NHEJ. In *A. nidulans*, by over-expressing the Rad51 homolog, uvsC, a modest increase in gene targeting frequencies were observed (from 5.9% to 26.9% and from 9.0% to 17.3%), not comparable to the effect observed by disabling NHEJ [59]. Furthermore both the thickness of the mycelia and conidiation was affected by the uvsC overexpression. While less severe, disabling of NHEJ is not without side-effects. While morphology is generally unchanged, several Aspergilli, including *A. oryzae* [60], *A. luchuensis* [61], *A. aculeatus* [62] and *A. fumigatus* [58] show increased sensitivity to methane methyl sulfonate (MMS), which is believed to stall the replication fork. Furthermore the Ku proteins, besides their involvement in NHEJ is also involved in telomeric maintenance, and the lack of Ku protein can cause telomeric shortening or sometimes lengthening in many organism including *S. cerevisiae* [41], *A. nidulans* [63], plants and mammals [42,64].

Genome editing technologies

An alternative to enhance gene targeting by manipulating the DNA repair pathways is instead to stimulate DNA repair at the desired location by causing a DSB at that locus. There are several gene editing technologies available, but in common for all of them are that they are all based on inducing a DSB at a specific location and then either cause small indel formation due to error-prone NHEJ or stimulate HR by addition of a homologous repair template. The creation of a DSB at a specific location is an incredible powerful tool, since then the cell's DNA repair can be harnessed to facilitate various modifications in the genome. This allows for modifications in organisms otherwise not amenable to genetic engineering. If a DSB is repaired by MMEJ or NHEJ, there is the chance for small insertions or deletions (indels) at the sites, especially if the causing agent is still present, since then only a mutation can stop repeated cutting, and then the function of the gene in question is likely to be disrupted by frameshift mutations. Alternatively by providing a homologous DNA template to stimulate repair by HR, it is possible to do targeted deletion, integrations or point mutations. A third approach is based on utilizing the MMEJ for doing targeted integrations [65]. Much of the motivation for developing genome editing tools comes from the desire to modify mammalian systems, which is notoriously difficult to engineer, but can provide valuable disease models, production host for various pharmaceutical or plants which can be of great agricultural or

commercial interests. In filamentous fungi the situation is different, since genetic engineering is feasible, but depending on various factors, can be time consuming. However depending on how resource intensive the technology in question is, it has the potential to be interesting for the possibility to make it easier to manipulate organisms as well.

The most common gene editing technologies include mega-nucleases, zinc-finger nucleases (ZFN), transcription activator-like effector nucleases (TALENs) and recently, clustered regularly interspaced palindromic repeats/ CRISPR-associated protein (CRISPR-Cas9).

Meganucleases

Meganucleases are a type of nucleases, characterized by their very long , 12-40 bp recognition sites [66]. An example of some of these meganucleases are the homing endonucleases, which are very specific nucleases, part of a mobile genetic element, encoded in introns or inteins, to facilitate the spread of that element to a specific targeting site by gene conversion [67]. A commonly used homing endonuclease is I-SceI, which originates from an intron in *S. cerevisiae* mitochondrial DNA[68]. While the long recognition sites, makes the meganuclease extremely specific, it is also one of the main disadvantages since the chances are that such a recognition site is present at the desired locus is next to none. As such many applications of meganucleases are based on pre-engineering a recognition site into the locus [69,70], and as such, while useful, its application for genome editing is limited.

Zinc-finger nucleases

Zinc-finger nucleases are fusion proteins, combining zinc-finger DNA binding domains with the cleavage domain of the FokI endonuclease [71]. Unusual for the FokI endonuclease is that the cleavage domain is separate from the DNA binding domain [72] and while it may exist in solution as a monomer, DNA cleavage only happens upon dimerization [73]. Further work has been done to promote hetero-dimerization over homo-dimerization [74,75], allowing for more specific cleavages by combing two different DNA binding domains. The underlying concept behind ZFN, is the Cys_2His_2 Zinc finger protein, where each zinc-finger recognizes three base pairs, and which can be assembled in a modular fashion [71,76]. For each monomer of FokI, generally three to six zinc-finger domains can be assembled, and then combined with a different monomer, for a hetero-dimer recognizing up to 36 bp [76]. The concept is illustrated in Figure 1.3A. Each zinc-finger motif consist of 30 amino acids with seven residues conferring the specificity for a certain triplet [77]. Despite the seemingly straightforward and modular design, several challenges are involved in the

design of functional ZFNs. One issues is that the sequence specificity and affinity are context dependent based on the neighbors ZFs, and that some configurations can lead to overlap in the targeting site [77] A main disadvantage of ZFN is the effort required to design and validate each ZFN, making it a costly process, preventing widespread use [78].

Transcription activator-like effector nucleases

TALENs are like ZFNs, hybrid proteins combining the FokI cleavage domain, with the transcription activator-like effectors (TALE) domains, from the plant pathogen *Xanthomonas* spp. In nature the TALE proteins are secreted to the plant, where they modulate expression of genes which supports bacterial virulence [79]. The DNA binding domain of the TALE proteins consists of small repeat region of most commonly 34 amino acids, however repeats from 30 to 42 amino acids exist [79], however the 12^{th} and 13^{th} are said to be hypervariable. These repeated regions are assembled in a modular fashions, and each repeat, depending on the hypervariable residues facilitate binding to a specific nucleotide, with hypervariable regions existing which allows for binding to any nucleotide or a subset of nucleotides. As such a consecutive array of these repeats facilitates binding to a consecutive DNA sequence, as illustrated in Figure 1.3B. Once the relationship between hypervariable regions and DNA binding was figured out [80,81], these binding domains, like ZFNs, have been fused to the FokI cleavage domain to create hybrid nucleases [82,83]. While TALENs are significantly easier/cheaper to engineer compared to ZFNs, the highly repetitive nature of the proteins presents a challenge since many cloning methods does not deal well with repetitive DNA.

Clustered regular interspaced palindromic repeats

CRISPR-Cas9 is the newest system for genome editing, but has rapidly gained popularity due to its simplicity. The genome editing tool originates from an adaptive immune defense system found in approximately 40% of bacteria and 90% of archaea [84]. The name refers to small DNA repeats, separated by spacer DNA and these structures were first discovered in 1987 in Japan in *Escheria coli* [85] and later in other prokaryotes and archaea [86], however their function was unknown. The function as an adaptive immune system was proposed in 2005 with the discovery that most of the spacer sequences was of plasmid or viral origin [87–89] together with the fact that these loci are

transcribed [90] and the term CRISPR was coined. A number of CRISPR associated (*cas*) genes was identified, including proteins with nucleases helicases domains. There are several types of CRISPR immune systems, the type used for genome editing and thus the type described in this section is a type II CRISPR immune system, but type I and III also exist [91]. Functionally the type II system works by uptake of invading DNA, which is then incorporated as spacers into these CRISPR arrays between direct repeats [91,92]. The arrays are then transcribed and matured into crispr RNA (crRNA), which consist of a repeat and a spacer. Another RNA, called a trans-activating CRISPR RNA (tracRNA), forms a duplex by basepairing with the repeat part of the crRNA, called a guide RNA (gRNA) [91,92]. Cas9 is a RNA-guided ribonuclease, and when it forms a complex with the gRNA, it can bind to DNA matching the spacer sequence, if the target DNA is followed by a protospacer adjacent motive (PAM), for which the specific sequence is a specific homolog of Cas9, and cleaves the target DNA [93]. This PAM sequence protects the cells from autoimmunity, ensuring that the CRISPR arrays will not be cut by the Cas proteins.

CRISPR for genome editing

The ability to target a specific sequence for cleavage by basepairing is incredible appealing, greatly simplifying the design compared to ZFNs and TALENs. Specifically for the Streptococcus pyogenes Cas9, which was the first one to be adapted to genome editing [94–98], the protospacer is 20 bp, which at the genomic locus should be followed by a PAM sequence consisting of NGG (NAG works as well, but with lower efficiencies [99]). As such, by exchanging only 20 bp, the Cas9 can be reprogrammed to a new target, being far easier and a lot cheaper compared to other genome editing technologies.

For genome editing the system has been simplified making a single chimeric gRNA, connecting the crRNA with the tracRNAs with a hairpin loop [93], making it only a two component system rather than three, with a single guide RNA (sgRNA) and the Cas9 nuclease, illustrated in Figure 1.3C. While the *S. pyogenes* Cas9 was the first to be used for genome editing, other variants of the system have similarly been used. Some of these systems have different PAM usage, and examples includes CRISPR-Cas systems from *Streptococcus thermophilus*, *Neisseria meningitidis* and *Staphylococcus aureus* [100,101], with the last one being particular interesting due to the Cas protein being about

25% smaller than that of the *S. pyogenes* Cas9, due to size limitation in some vector systems.

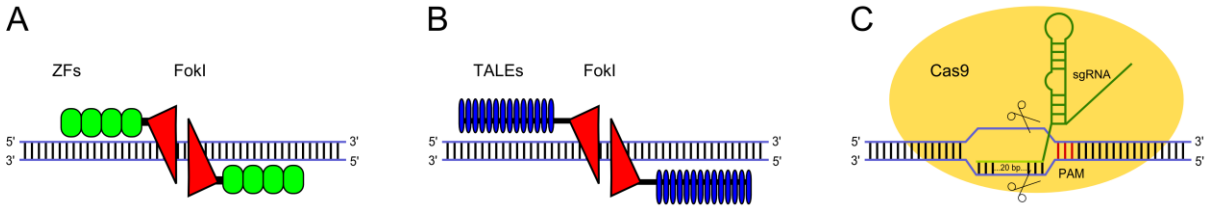

Figure 1.3 Three types of programmable genome editing technologies. A) ZFs assembled in a modular fashion and fused to FokI. Each ZF binds a DNA triplet. B) TALEs assembled in a modular fashion and fused to FokI. Each TALE recognize one base. C) CRISPR-Cas9. 20 bases of the sgRNA binds to the genomic target by basepairing. The remaining part of the sgRNA is in complex with the Cas9 protein, directing cleavage of the target DNA.

Off-targeting

Off-targeting is an important concern for all types of genome editing technologies, since the introduction of a DSB in the wrong position can lead to undesired mutations. There are two aspects to consider regarding off-targeting. One aspect, which technically is lack of the necessary specificity, rather than off-targeting, is the length of the targeting sequence, since for short targeting sequences, especially in large genomes the chance of the same sequence occurring more than once, can be high. The "real" off-targeting issues stems from the fact that basically all of the genome editing technologies have issues with the respective nuclease that sometimes cleaves sequences, while similar, is not an exact match to the recognition sequence. Lots of work has been done to limit off-targeting in the various genome editing technologies.

Both ZFNs and TALENs are assembled in a modular fashion and both relies on the FokI endonuclease. That have also dictated the two main focus areas for increasing specificity. One focus has been on assembling enough modules, ZFs or TALEs, increasing the recognition site to get sufficient specificity, which is also related to genome size. The other common factor is the FokI endonuclease, which have been engineered to be an obligate heterodimer, essentially doubling the recognition site [75,102]. Similar strategies have also been applied to the CRISPR-Cas9. The Cas9 nuclease have two domains each responsible for cleaving one of the DNA strands [93]. One strategy is based on inactivating one of these domains, turning the nuclease into a nickase, in combination with two sgRNAs with will cause two nicks on opposite strand which can stimulate gene targeting similar to a single DSB [103]. Another strategy, similar to that of ZFNs and TALENs is based on inactivating both cleavage domains of Cas9 and instead fusing it to FokI, so that basepairing of two sgRNAs in close proximity is necessary to cause cleavage by FokI [104].

However, when it comes to the off-targeting issues, the main focus has been on predicting the parameters for when it happens. Especially ZFNs is tricky since the individual ZFs have affinity for more than one triplet [105], in addition to the effects from neighboring ZFs. To optimize ZFNs the focus have been on designing various platforms for efficient selection of ZFNs [106–108]. For TALENs off-targeting is likewise an issue, but the extent of it is still largely undiscovered [109]. For the CRISPR-Cas9 lots of effort has been put into figuring out the conditions triggering off-target cleavage, and while the specificity of CRISPR-Cas9 might be less than that of TALENs simply due to the fact that CRISPR-Cas9 cleavage is based on a 22bp recognition sites, while that of a TALEN can be significantly longer 24-40bp [110], since targeting is based on basepairing potential off-targets can more easily be predicted. Considering CRISPR-Cas9 is based on an immune system, it might make sense that some degree of mismatching is tolerated to limit pathogens escaping by mutating. Two studies found that while mismatches generally are not tolerated in the PAM proximal end, while up to three mismatched in the PAM distal end could be tolerated [99,111]. Another studies found that 2bp mismatches caused no or very little off-targeting [112], and studies in *S. cerevisiae* or CHO cells found that off-targeting was not an issue [113,114]. Overall this shows that there is still a lot to figure out regarding off-targeting with CRISPR-Cas.

Genome editing in Aspergillus

While the first reports on the use of CRISPR-Cas9 for genome editing came out in the start of 2013, the first reports on CRISPR-Cas9 being used for the genetic engineering of filamentous fungi did not show up until mid 2015, and one of those is presented in chapter three. Furthermore until CRISPR-Cas9, genome editing technologies have not been used in filamentous fungi, except for a single report on the use of TALENs in *Pyricularia oryzae* [115]. Compared to previous genome editing technologies the cost and effort of using CRISPR-Cas9 is low enough that it is being applied to organisms otherwise amendable to genetic engineering with great results, such as *S. cerevisiae* [113,116,117], while it has greatly transformed genetic engineering of plant and mammalian systems [91]. This shows great promise that it will also turn out to be a valuable tool for genetic engineering of filamentous fungi.

References

1. Mäkelä MR, Donofrio N, Vries RP De. Plant biomass degradation by fungi. 2014;72: 2–9.

2. Lubertozzi D, Keasling JD. Developing Aspergillus as a host for heterologous expression [Internet]. Biotechnology Advances. Elsevier Inc.; 2009. pp. 53–75. doi:10.1016/j.biotechadv.2008.09.001

3. Pontecorvo G. The genetic of Aspergillus nidulans. 1953.

4. Morris NR. Mitotic mutants of Aspergillus nidulans. Genet Res. 1975;26: 237–254. doi:10.1017/S0016672300016049

5. Machida M, Yamada O, Gomi K. Genomics of Aspergillus oryzae: Learning from the History of Koji Mold and Exploration of Its Future. DNA Res. 2008;15: 173–183. doi:10.1093/dnares/dsn020

6. Dagenais TRT, Keller NP. Pathogenesis of Aspergillus fumigatus in Invasive Aspergillosis. Clin Microbiol Rev. 2009;22: 447–465. doi:10.1128/CMR.00055-08

7. Machida M, Asai K, Sano M, Tanaka T, Kumagai T, Terai G, et al. Genome sequencing and analysis of Aspergillus oryzae. Nature. 2005;438: 1157–1161. doi:10.1038/nature04300

8. Galagan JE, Calvo SE, Cuomo C, Ma L-J, Wortman JR, Batzoglou S, et al. Sequencing of Aspergillus nidulans and comparative analysis with A. fumigatus and A. oryzae. Nature. 2005;438: 1105–1115. doi:10.1038/nature04341

9. Nierman WC, Pain A, Anderson MJ, Wortman JR, Kim HS, Arroyo J, et al. Genomic sequence of the pathogenic and allergenic filamentous fungus Aspergillus fumigatus. Nature. 2005;438: 1151–1156. doi:10.1038/nature04332

10. Pel HJ, de Winde JH, Archer DB, Dyer PS, Hofmann G, Schaap PJ, et al. Genome sequencing and analysis of the versatile cell factory Aspergillus niger CBS 513.88. Nat Biotechnol. 2007;25: 221–231. doi:10.1038/nbt1282

11. Case ME, Schweizer M, Kushner SR, Giles NH. Efficient transformation of Neurospora crassa by utilizing hybrid plasmid DNA. Proc Natl Acad Sci. 1979;76: 5259–5263. doi:10.1073/pnas.76.10.5259

12. Tilburn J, Scazzocchio C, Taylor GG, Zabicky-Zissman JH, Lockington RA, Davies RW. Transformation by integration in Aspergillus nidulans. Gene. 1983;26: 205–221. doi:10.1016/0378-1119(83)90191-9

13. Michael Ward, Katherine H. Kodama and LJW. Transformation of Aspergillus awamori by

Electroporation and A . niger. Exp Mycol. 1989;13: 289–293.

14. Herzog RW, Daniell H, Singh NK, Lemke P a. A comparative study on the transformation of Aspergillus nidulans by microprojectile bombardment of conidia and a more conventional procedure using protoplasts treated with polyethyleneglycol. Appl Microbiol Biotechnol. 1996;45: 333–337. doi:10.1007/s002530050692

15. Gouka RJ, Gerk C, Hooykaas PJ, Bundock P, Musters W, Verrips CT, et al. Transformation of Aspergillus awamori by Agrobacterium tumefaciens-mediated homologous recombination. Nat Biotechnol. 1999;17: 598–601. doi:10.1038/9915

16. Gems D, Johnstone IL, Clutterbuck AJ. An autonomously replicating plasmid transforms Aspergillus nidulans at high frequency. Gene. 1991;98: 61–67. doi:10.1016/0378-1119(91)90104-J

17. Cullen D, Leong SA, Wilson LJ, Henner DJ. Transformation of Aspergillus nidulans with the hygromycin-resistance gene, hph. Gene. 1987;57: 21–26. doi:10.1016/0378-1119(87)90172-7

18. Punt PJ, Oliver RP, Dingemanse MA, Pouwels PH, van den Hondel CAMJJ. Transformation of Aspergillus based on the hygromycin B resistance marker from Escherichia coli. Gene. 1987;56: 117–124. doi:10.1016/0378-1119(87)90164-8

19. Austin B, Hall RM, Tyler BM. Optimized vectors and selection for transformation of Neurospora crassa and Aspergillus nidulans to bleomycin and phleomycin resistance. Gene. 1990;93: 157–162. doi:10.1016/0378-1119(90)90152-H

20. Ward M, Wilson LJ, Carmona CL, Turner G. The oliC3 gene of Aspergillus niger: isolation, sequence and use as a selectable marker for transformation. Curr Genet. 1988;14: 37–42. doi:10.1007/BF00405851

21. Ballance DJ, Buxton FP, Turner G. Tranformation of Aspergillus nidulans by the orotidine-5´-phosphate decarboxylase gene of Neuraspora crassa. Biochmical Biophys Res Commun. 1983;112: 284–289.

22. Hynes MJ, Pateman JAJ. The genetic analysis of regulation of amidase synthesis inAspergillus nidulans. MGG Mol Gen Genet. 1970;108: 97–106. doi:10.1007/BF02430516

23. Kelly JM, Hynes MJ. Transformation of Aspergillus niger by the amdS gene of Aspergillus nidulans. EMBO J. 1985;4: 475–479.

24. Devchand M, Gwynne DI. Expression of heterologous proteins in Aspergillus. J Biotechnol. 1991;17: 3–9. doi:10.1016/0168-1656(91)90022-N

25. Fleißner A, Dersch P. Expression and export: recombinant protein production systems for Aspergillus. Appl Microbiol Biotechnol. 2010;87: 1255–1270. doi:10.1007/s00253-010-2672-6

26. Vogt K, Bhabhra R, Rhodes JC, Askew DS. Doxycycline-regulated gene expression in the opportunistic fungal pathogen Aspergillus fumigatus. BMC Microbiol. 2005;5: 1. doi:10.1186/1471-2180-5-1

27. Meyer V, Wanka F, van Gent J, Arentshorst M, van den Hondel C a MJJ, Ram AFJ. Fungal Gene Expression on Demand: an Inducible, Tunable, and Metabolism-Independent Expression System for Aspergillus niger. Appl Environ Microbiol. 2011;77: 2975–2983. doi:10.1128/AEM.02740-10

28. Punt PJ, Dingemanse MA, Kuyvenhoven A, Soede RDM, Pouwels PH, Vandenhondel CAMJJ. Functional Elements in the Promoter Region of the Aspergillus-Nidulans Gpda Gene Encoding Glyceraldehyde-3-Phosphate Dehydrogenase. Gene. 1990;93: 101–109. Available: <Go to ISI>://A1990EC87400015

29. Kitamoto N, Matsui J, Kawai Y, Kato A, Yoshino S, Ohmiya K, et al. Utilization of the TEF1-?? gene (TEF1) promoter for expression of polygalacturonase genes, pgaA and pgaB, in Aspergillus oryzae. Appl Microbiol Biotechnol. 1998;50: 85–92. doi:10.1007/s002530051260

30. Geisberg J V, Moqtaderi Z, Fan X, Ozsolak F, Struhl K. Global analysis of mRNA isoform half-lives reveals stabilizing and destabilizing elements in yeast. Cell. 2014;156: 812–24. doi:10.1016/j.cell.2013.12.026

31. Curran K a., Morse NJ, Markham K a., Wagman AM, Gupta A, Alper HS. Short Synthetic Terminators for Improved Heterologous Gene Expression in Yeast. ACS Synth Biol. 2015;4: 824–832. doi:10.1021/sb5003357

32. Ottaviani A, Gilson E, Magdinier F. Telomeric position effect: from the yeast paradigm to human pathologies? Biochimie. 2008;90: 93–107. doi:10.1016/j.biochi.2007.07.022

33. Flagfeldt DB, Siewers V, Huang L, Nielsen J. Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae. Yeast. 2009;26: 545–551. doi:10.1002/yea.1705

34. Mikkelsen MD, Buron LD, Salomonsen B, Olsen CE, Hansen BG, Mortensen UH, et al. Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. Metab Eng. 2012; doi:10.1016/j.ymben.2012.01.006

35. Hansen BG, Salomonsen B, Nielsen MT, Nielsen JB, Hansen NB, Nielsen KF, et al. Versatile enzyme expression and characterization system for Aspergillus nidulans, with the Penicillium brevicompactum polyketide synthase gene from the mycophenolic acid gene cluster as a test case. Appl Environ Microbiol. 2011;77: 3044–3051. doi:10.1128/AEM.01768-10

36. Verdoes JC, Punt PJ, van den Hondel CAMJJ. Molecular genetic strain improvement for the overproduction of fungal proteins by Filamentous fungi. Appl Microbiol Biotechnol. 1995;43: 195–

205. doi:10.1007/s002530050390

37. Symington LS, Gautier J. Double-Strand Break End Resection and Repair Pathway Choice. Annu Rev Genet. 2011;45: 247–271. doi:10.1146/annurev-genet-110410-132435

38. Doherty AJ, Jackson SP, Weller GR. Identification of bacterial homologues of the Ku DNA repair proteins. FEBS Lett. 2001;500: 186–188. doi:10.1016/S0014-5793(01)02589-3

39. Weller GR, Kysela B, Roy R, Tonkin LM, Scanlan E, Della M, et al. Identification of a DNA nonhomologous end-joining complex in bacteria. Science. 2002;297: 1686–1689. doi:10.1126/science.1074584

40. Doherty AJ, Jackson SP, Weller GR. Identification of bacterial homologues of the Ku DNA repair proteins [2]. FEBS Letters. 2001. pp. 186–188. doi:10.1016/S0014-5793(01)02589-3

41. Boulton SJ, Jackson SP. Identification of a Saccharomyces cerevisiae Ku80 homologue: roles in DNA double strand break rejoining and in telomeric maintenance. Nucleic Acids Res. 1996;24: 4639–4648. doi:10.1093/nar/24.23.4639

42. Fisher TS, Zakian V a. Ku: A multifunctional protein involved in telomere maintenance. DNA Repair. 2005. pp. 1215–1226. doi:10.1016/j.dnarep.2005.04.021

43. Wu D, Topper LM, Wilson TE. Recruitment and dissociation of nonhomologous end joining proteins at a DNA double-strand break in Saccharomyces cerevisiae. Genetics. 2008;178: 1237–1249. doi:10.1534/genetics.107.083535

44. Daley JM, Palmbos PL, Wu D, Wilson TE. Nonhomologous end joining in yeast. Annu Rev Genet. 2005;39: 431–451. doi:10.1146/annurev.genet.39.073003.113340

45. Pardo B, Gómez-González B, Aguilera a. DNA double-strand break repair: How to fix a broken relationship. Cell Mol Life Sci. 2009;66: 1039–1056. doi:10.1007/s00018-009-8740-3

46. Krogh BO, Symington LS. Recombination proteins in yeast. Annu Rev Genet. 2004;38: 233–271. doi:10.1146/annurev.genet.38.072902.091500

47. Liu J, Renault L, Veaute X, Fabre F, Stahlberg H, Heyer W-D. Rad51 paralogues Rad55–Rad57 balance the antirecombinase Srs2 in Rad51 filament formation. Nature. Nature Publishing Group; 2011;479: 245–248. doi:10.1038/nature10522

48. Mortensen UH, Lisby M, Rothstein R. Rad52. DNA Repair (Amst). 2009;8: 1068–1076. doi:10.1016/j.dnarep.2009.04.007

49. Mazin A V., Mazina OM, Bugreev D V., Rossi MJ. Rad54, the motor of homologous recombination.

DNA Repair (Amst). 2010;9: 286–302. doi:10.1016/j.dnarep.2009.12.006

50. Mehta A, Haber JE. Sources of DNA Double-Strand Breaks and Models of Recombinational DNA Repair. Cold Spring Harb Perspect Biol. 2014; doi:10.1101/cshperspect.a016428

51. Langston LD, Symington LS. Gene targeting in yeast is initiated by two independent strand invasions. Proc Natl Acad Sci U S A. 2004;101: 15392–15397. doi:10.1073/pnas.0403748101

52. Deng SK, Gibb B, de Almeida MJ, Greene EC, Symington LS. RPA antagonizes microhomology-mediated repair of DNA double-strand breaks. Nat Struct Mol Biol. Nature Publishing Group; 2014;21: 405–12. doi:10.1038/nsmb.2786

53. Shrivastav M, De Haro LP, Nickoloff JA. Regulation of DNA double-strand break repair pathway choice. Cell Res. 2008;18: 134–147. doi:10.1038/cr.2007.111

54. Hua SB, Qiu M, Chan E, Zhu L, Luo Y. Minimum length of sequence homology required for in vivo cloning by homologous recombination in yeast. Plasmid. 1997;38: 91–96. doi:10.1006/plas.1997.1305

55. Nayak T. A Versatile and Efficient Gene-Targeting System for Aspergillus nidulans. Genetics. 2005;172: 1557–1566. doi:10.1534/genetics.105.052563

56. Takahashi T, Masuda T, Koyama Y. Enhanced gene targeting frequency in ku70 and ku80 disruption mutants of Aspergillus sojae and Aspergillus oryzae. Mol Genet Genomics. 2006;275: 460–470. doi:10.1007/s00438-006-0104-1

57. Krappmann S, Sasse C, Braus GH. Gene Targeting in Aspergillus fumigatus by Homologous Recombination Is Facilitated in a Nonhomologous End- Joining-Deficient Genetic Background Gene Targeting in Aspergillus fumigatus by Homologous Recombination Is Facilitated in a Nonhomologous End- Join. Eukaryot Cell. 2006; doi:10.1128/EC.5.1.212

58. da Silva Ferreira ME, Kress MRVZ, Savoldi M, Goldman MHS, Härtl A, Heinekamp T, et al. The akuBKU80 Mutant Deficient for Nonhomologous End Joining Is a Powerful Tool for Analyzing Pathogenicity in Aspergillus fumigatus The akuB KU80 Mutant Deficient for Nonhomologous End Joining Is a Powerful Tool for Analyzing Pathogenicity in Aspergillus. Eukaryot Cell. 2006; doi:10.1128/EC.5.1.207

59. Natsume T, Egusa M, Kodama M, Johnson R, Itoh T, Itoh Y. An appropriate increase in the transcription of Aspergillus nidulans uvsC improved gene targeting efficiency. Biosci Biotechnol Biochem. 2004;68: 1649–56. doi:10.1271/bbb.68.1649

60. Mizutani O, Kudo Y, Saito A, Matsuura T, Inoue H, Abe K, et al. A defect of LigD (human Lig4 homolog) for nonhomologous end joining significantly improves efficiency of gene-targeting in Aspergillus oryzae. Fungal Genet Biol. 2008;45: 878–889. doi:10.1016/j.fgb.2007.12.010

61. Takahashi T, Mizutani O, Shiraishi Y, Yamada O. Development of an efficient gene-targeting system in Aspergillus luchuensis by deletion of the non-homologous end joining system. J Biosci Bioeng. The Society for Biotechnology, Japan; 2011;112: 529–534. doi:10.1016/j.jbiosc.2011.08.007

62. Tani S, Tsuji A, Kunitake E, Sumitani J, Kawaguchi T. Reversible impairment of the ku80 gene by a recyclable marker in Aspergillus aculeatus. AMB Express. AMB Express; 2013;3: 4. doi:10.1186/2191-0855-3-4

63. Nielsen JB, Nielsen ML, Mortensen UH. Transient disruption of non-homologous end-joining facilitates targeted genome manipulations in the filamentous fungus Aspergillus nidulans. Fungal Genet Biol. 2008;45: 165–170. doi:10.1016/j.fgb.2007.07.003

64. Hande MP. DNA repair factors and telomere-chromosome integrity in mammalian cells. Cytogenetic and Genome Research. 2004. pp. 116–122. doi:10.1159/000077475

65. Nakade S, Tsubota T, Sakane Y, Kume S, Sakamoto N, Obara M, et al. Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. Nat Commun. Nature Publishing Group; 2014;5: 5560. doi:10.1038/ncomms6560

66. Belfort M, Roberts RJ. Homing endonucleases: keeping the house in order. Nucleic Acids Res. 1997;25: 3379–88. Available: http://www.ncbi.nlm.nih.gov/pubmed/9254693

67. Stoddard BL. Homing endonuclease structure and function. Q Rev Biophys. 2005;38: 49–95. doi:10.1017/S0033583505004063

68. Jurica MS, Stoddard BL. Homing endonucleases: Structure, function and evolution [Internet]. Cellular and Molecular Life Sciences. 1999. pp. 1304–1326. doi:10.1007/s000180050372

69. Storici F, Resnick M a. The Delitto Perfetto Approach to In Vivo Site-Directed Mutagenesis and Chromosome Rearrangements with Synthetic Oligonucleotides in Yeast. Methods in Enzymology. 2006. pp. 329–345. doi:10.1016/S0076-6879(05)09019-1

70. Alexander WG, Doering DT, Hittinger CT. High-Efficiency Genome Editing and Allele Replacement in Prototrophic and Wild Strains of Saccharomyces. Genetics. 2014;198: 859–866. doi:10.1534/genetics.114.170118

71. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A. 1996;93: 1156–1160. doi:10.1073/pnas.93.3.1156

72. Li L, Wu LP, Chandrasegaran S. Functional domains in Fok I restriction endonuclease. Proc Natl Acad Sci U S A. 1992;89: 4275–4279. doi:10.1073/pnas.89.10.4275

73. Bitinaite J, Wah D a, Aggarwal a K, Schildkraut I. FokI dimerization is required for DNA cleavage. Proc Natl Acad Sci U S A. 1998;95: 10570–10575. doi:10.1073/pnas.95.18.10570

74. Miller JC, Holmes MC, Wang J, Guschin DY, Lee Y-L, Rupniewski I, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol. 2007;25: 778–785. doi:10.1038/nbt1319

75. Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, et al. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. Nat Methods. 2011;8: 74–79. doi:10.1038/nmeth.1539

76. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. Nat Rev Genet. Nature Publishing Group; 2010;11: 636–646. doi:10.1038/nrg2842

77. Durai S, Mani M, Kandavelou K, Wu J, Porteus MH, Chandrasegaran S. Zinc finger nucleases: Custom-designed molecular scissors for genome engineering of plant and mammalian cells. Nucleic Acids Res. 2005;33: 5978–5990. doi:10.1093/nar/gki912

78. DeFrancesco L. Move over ZFNs. Nat Biotechnol. 2011;29: 681–4. doi:10.1038/nbt.1935

79. Boch J, Bonas U. Xanthomonas AvrBs3 family-type III effectors: discovery and function. Annu Rev Phytopathol. 2010;48: 419–436. doi:10.1146/annurev-phyto-080508-081936

80. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. Science. 2009;326: 1509–1512. doi:10.1126/science.1178811

81. Moscou MJ, Bogdanove AJ. A Simple Cipher Governs DNA Recognition by TAL Effectors. Science (80-). 2009;326: 1501–1501. doi:10.1126/science.1178817

82. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, et al. Targeting DNA double-strand breaks with TAL effector nucleases. Genetics. 2010;186: 756–761. doi:10.1534/genetics.110.120717

83. Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP, et al. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. Nucleic Acids Res. 2011;39: 359–372. doi:10.1093/nar/gkq704

84. Horvath P, Barrangou R. CRISPR/Cas, the Immune System of Bacteria and Archaea. Science (80-). 2010;327: 167–170. doi:10.1126/science.1179555

85. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J Bacteriol. 1987;169: 5429–5433.

86. Mojica FJM, Diez-Villasenor C, Soria E, Juez G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. Mol Microbiol. 2000;36: 244–246. doi:10.1046/j.1365-2958.2000.01838.x

87. Bolotin A, Quinquis B, Sorokin A, Dusko Ehrlich S. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology. 2005;151: 2551–2561. doi:10.1099/mic.0.28048-0

88. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol. 2005;60: 174–82. doi:10.1007/s00239-004-0046-3

89. Pourcel C, Salvignol G, Vergnaud G. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology. 2005;151: 653–663. doi:10.1099/mic.0.27437-0

90. Tang T-H, Bachellerie J-P, Rozhdestvensky T, Bortolin M-L, Huber H, Drungowski M, et al. Identification of 86 candidates for small non-messenger RNAs from the archaeon Archaeoglobus fulgidus. Proc Natl Acad Sci. 2002;99: 7536–7541. doi:10.1073/pnas.112047299

91. Doudna JA, Charpentier E. The new frontier of genome engineering with CRISPR-Cas9. Science (80-). 2014;346: 1258096–1258096. doi:10.1126/science.1258096

92. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol. 2014;32: 347–355. doi:10.1038/nbt.2842

93. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science (80-). 2012;337: 816–821. doi:10.1126/science.1225829

94. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. Elife. 2013;2: e00471. doi:10.7554/eLife.00471

95. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol. Nature Publishing Group; 2013;31: 227–229. doi:10.1038/nbt.2501

96. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339: 819–23. doi:10.1126/science.1231143

97. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-Guided Human Genome Engineering via Cas9. Science (80-). 2013;339: 823–826. doi:10.1126/science.1232033

98. Cho SW, Kim S, Kim JM, Kim J-S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol. Nature Publishing Group; 2013;31: 230–232. doi:10.1038/nbt.2507

99. Hsu PD, Scott D a, Weinstein J a, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol. 2013;31: 827–32. doi:10.1038/nbt.2647

100. Xu K, Ren C, Liu Z, Zhang T, Zhang T, Li D, et al. Efficient genome engineering in eukaryotes using Cas9 from Streptococcus thermophilus. Cell Mol Life Sci. 2015;72: 383–399. doi:10.1007/s00018-014-1679-z

101. Ran FA, Cong L, Yan WX, Scott D a., Gootenberg JS, Kriz AJ, et al. In vivo genome editing using Staphylococcus aureus Cas9. Nature. 2015;520: 186–191. doi:10.1038/nature14299

102. Miller JC, Holmes MC, Wang J, Guschin DY, Lee Y-L, Rupniewski I, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol. 2007;25: 778–785. doi:10.1038/nbt1319

103. Ran FA, Hsu PD, Lin C-Y, Gootenberg JS, Konermann S, Trevino AE, et al. Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. Cell. Elsevier; 2013;154: 1380–1389. doi:10.1016/j.cell.2013.08.021

104. Tsai SQ, Wyvekens N, Khayter C, Foden J a, Thapar V, Reyon D, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nat Biotechnol. 2014;32: 569–576. doi:10.1038/nbt.2908

105. Carroll D. Genome Engineering With Zinc-Finger Nucleases. Genetics. 2011;188: 773–782. doi:10.1534/genetics.111.131433

106. Maeder ML, Thibodeau-Beganny S, Sander JD, Voytas DF, Joung JK. Oligomerized pool engineering (OPEN): an “open-source” protocol for making customized zinc-finger arrays. Nat Protoc. 2009;4: 1471–1501. doi:10.1038/nprot.2009.98

107. Gonzalez B, Schwimmer LJ, Fuller RP, Ye Y, Asawapornmongkol L, Barbas CF. Modular system for the construction of zinc-finger libraries and proteins. Nat Protoc. Nature Publishing Group; 2010;5: 791–810. doi:10.1038/nprot.2010.34

108. Sander JD, Dahlborg EJ, Goodwin MJ, Cade L, Zhang F, Cifuentes D, et al. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). Nat Methods. 2011;8: 67–69. doi:10.1038/nmeth.1542

109. Li M, Suzuki K, Kim NY, Liu GH, Belmonte JCI. A cut above the rest: Targeted genome editing technologies in human pluripotent stem cells. J Biol Chem. 2014;289: 4594–4599. doi:10.1074/jbc.R113.488247

110. Gaj T, Gersbach C a, Barbas CF. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering [Internet]. Trends in Biotechnology. Elsevier Ltd; 2013. pp. 397–405. doi:10.1016/j.tibtech.2013.04.004

111. Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol. 2013;31: 833–838. doi:10.1038/nbt.2675

112. Kim D, Bae S, Park J, Kim E, Kim S, Yu HR, et al. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. Nat Methods. 2015;12: 237–43, 1 p following 243. doi:10.1038/nmeth.3284

113. Jakočiūnas T, Bonde I, Herrgård M, Harrison SJ, Kristensen M, Pedersen LE, et al. Multiplex metabolic pathway engineering using CRISPR/Cas9 in Saccharomyces cerevisiae. Metab Eng. 2015;28: 213–222. doi:10.1016/j.ymben.2015.01.008

114. Grav LM, Lee JS, Gerling S, Kallehauge TB, Hansen AH, Kol S, et al. One-step generation of triple knockout CHO cell lines using CRISPR/Cas9 and fluorescent enrichment. Biotechnol J. 2015;10: 1446–1456. doi:10.1002/biot.201500027

115. Arazoe T, Ogawa T, Miyoshi K, Yamato T, Ohsato S, Sakuma T, et al. Tailor-made TALEN system for highly efficient targeted gene replacement in the rice blast fungus. Biotechnol Bioeng. 2015;112: 1335–1342. doi:10.1002/bit.25559

116. Bao Z, Xiao H, Liang J, Zhang L, Xiong X, Sun N, et al. Homology-integrated CRISPR-Cas (HI-CRISPR) system for one-step multigene disruption in Saccharomyces cerevisiae. ACS Synth Biol. 2015;4: 585–94. doi:10.1021/sb500255k

117. Mans R, van Rossum HM, Wijsman M, Backx A, Kuijpers NG a., van den Broek M, et al. CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in Saccharomyces cerevisiae. FEMS Yeast Res. 2015;15: fov004–fov004. doi:10.1093/femsyr/fov004

Chapter 2: Intronic insertion as a tool for creating transient gene disruptions

Introduction

The ability to transiently disable a gene of choice can be a valuable tool for many purposes. Options include disabling of nutritional important genes to create auxotrophic selection markers, disabling of genes in the non-homologous end-joining (NHEJ) pathway to facilitate high frequencies of gene targeting or temporarily insertion of genes that you eventually would like to get rid of. Previous work in the field has been done to transiently disable *pyrG* in *A. niger* [1] for creating a strain auxotrophic for uridine, but resistant to 5-FOA. Likewise, *nkuA* in *A. nidulans* has been transiently disabled for boosting gene targeting frequencies [2]. These transient disruptions help bridging the needs between strains having traits suitable for genetic engineering with that of strains being suitable in an industrial setting. Unfortunately, the two approaches are mutually exclusive since both are taking advantages of the selectable and counter-selectable properties of *pyrG* as a nutritional selection marker and direct repeat recombination for excision. An alternative to relying on direct repeat recombination, is to take advantage of recombinases, such as the cre-lox, from the P1 bacteriophage [3], the FLP-FRT from the *S. cerevisiae* 2µ plasmid [4,5] or the β-six first isolated from *S. pyogenes* [6]. These recombinases, despite their very different origin, have each been shown to work in a wide range of species, including outside the kingdom in which they were isolated [7]. In general, these systems are two component systems consisting of a recombinase and a recognition sequence, and works by the recombinase catalyzing DNA combination between two of the recognition sites. The exact effect depends both on the specific system and on the orientation of the sites. For example for the cre-lox systems, if two loxP sites are in direct orientation and the cre recombinase is present, it will catalyze an excision event between the loxP sites, while if the loxP sites are inverted in regards to each other, the cre recombinase will facilitate an inversion on the DNA between the two sites [8].

A disadvantage to recombinase facilitated recombination over direct repeat combination is that while direct repeat recombination can be designed to leave no scar, recombinase facilitated recombination will leave one recognition site at the locus. However, unlike direct repeat recombination, the recombination frequencies are high enough to allow for successful

recombination events to be identified without the need for selection. Another useful feature is that it is possible to do multiple recombinase excision events at once. Therefore, to circumvent the issues with leaving a scar behind, we hypothesized that it will be possible to temporarily disrupt the function of a gene by inserting a marker gene into the gene of choice. We expect that the dramatic increase in intron size will interfere with intron splicing and lead to aberrant transcripts and thus no functional protein. To avoid impairing gene function permanently, by flanking the marker with loxP sites and target it to an intron, later excision by cre mediated recombination, leaving only the loxP scar in the intron, the function can be restored since the loxP scar is most likely small enough not to disrupt splicing of the intron. The whole concept is illustrated in Figure 2.1. Similar concepts have been utilized in both *S. cerevisiae* [9] and mouse models [10]. Despite the potential applications of transient knockouts based on intron-blocking, such a concept has not been applied to filamentous fungi before.

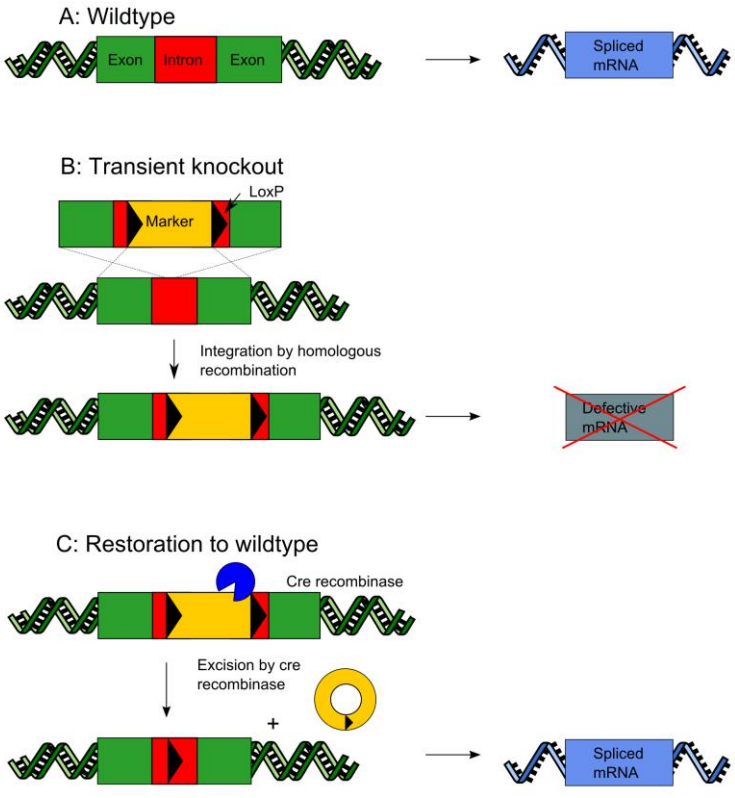

Figure 2.1 Concept for intron based transient gene disrupting. A) Wildtype situation, introns on the DNA levels, get spliced out on the RNA levels. B) A marker gene flanked by loxP sites and inserted into the intron by HR, and this blocks successful splicing and thus also translation. C) Exposure to the cre recombinase will excise the inserted marker gene from the intron, leaving only a single loxP scar. If placed correctly this will lead to successful splicing, similar to the wildtype intron.

Results

Proof of concept, transient disruption of yA

To test if the idea was feasible, the first objective was to make a transient disruption of the *yA* pigment gene in *A. nidulans* as a proof of concept. Donor, acceptor and splice sites were identified, based on consensus sequences found by Kupfer et al [11], since we hypothesized that if any of these elements were disrupted by the loxP scar, which would eventually be left behind, it would not be tolerated, regarding intron splicing. Two DNA constructs, pyA-i2::pyrG and pyA-i3::pyrG were made containing a *pyrG* selection marker, from *A. fumigatus*, flanked by two loxP sites in direct orientation, flanked by two targeting sequences targeting the vector to either the second or the third intron in *yA*. The recipient strain, NID1 (Figure 2.2, panel A) was defective for NHEJ and auxotrophic for arginine, uridine and uracil, and all transformed colonies from both constructs (Figure 2, panel B and D), turned yellow indicating that *yA* has been successfully disrupted. Two colonies of each, (NID2010 and NID2011 disrupted in the second *yA* intron and NID2012 and NID2013 disrupted in third *yA* intron) and were purify streaked, subjected to protoplastation and transformed with pCCM1, an AMA1 vector containing the cre recombinase under control of the *A. nidulans gpdA* promoter and the *argB* selection marker. After transformation, the resulting colonies from all four strains were green (Figure 2, panel C and E). However a minority of the colonies displayed few specks of yellow, demonstrating that a loxP scar through careful placing could be tolerated in the intron.

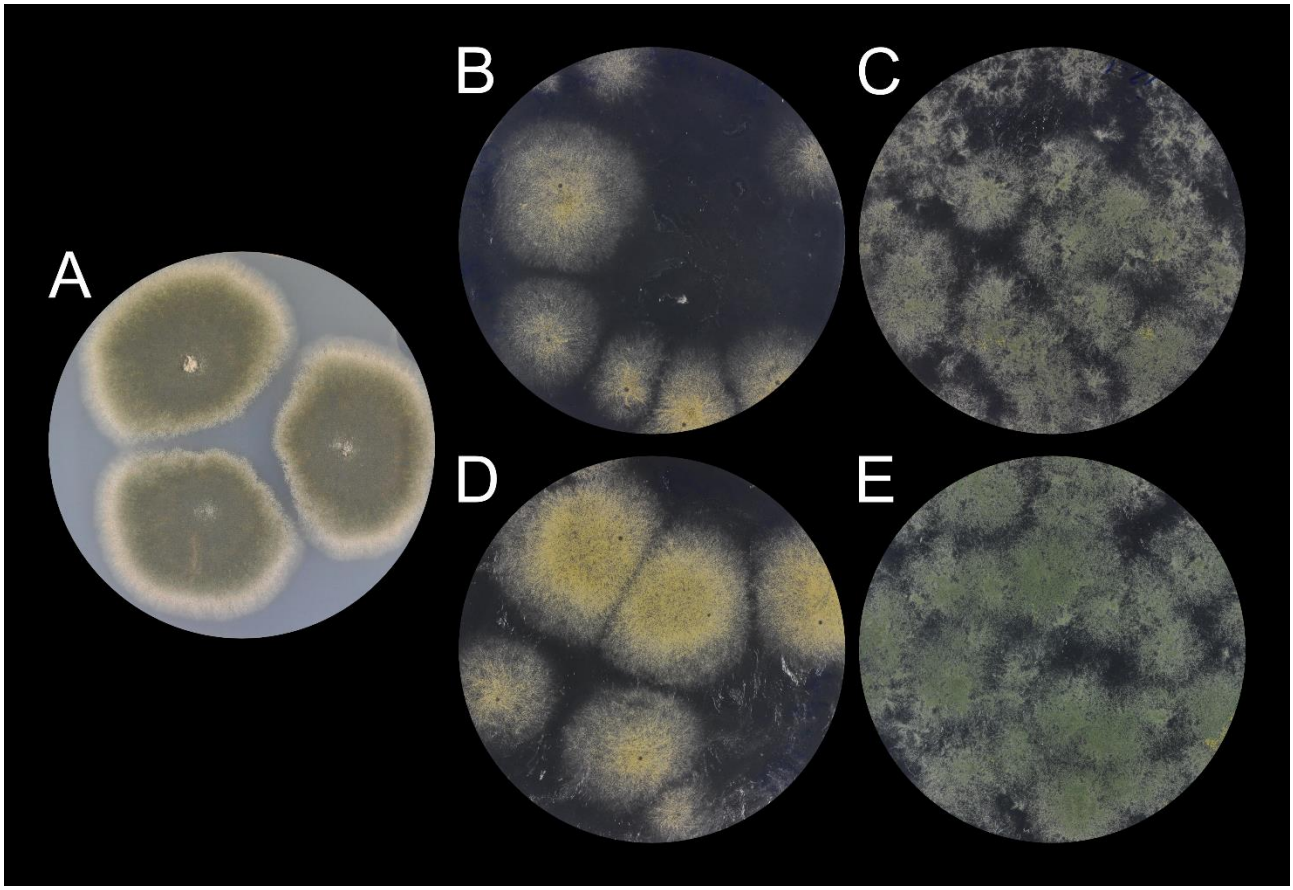

Figure 2.2 Proof of concept for transient gene disruption. A) Stab of NID1. B) Transformation plate of NID1 transformed with pyA-i2::pyrG, disrupting *yA* in the second intron, resulting in a yellow-spored phenotype. C) Transformation of one of the yellow colonies from panel B, with a vector containing the gene encoding the cre-recombinase, restoring the green-spored phenotype. D) Similar as B, but the vector facilitate a disruption of the third *yA* intron instead. E) Transformation of one of the yellow colonies from panel D, with a vector containing the gene encoding the cre-recombinase, restoring the green-spored phenotype.

To ensure that the loxP scar left in the intron did not affect expression levels of the transiently disrupted gene, qRT-PCR was performed comparing expression of wildtype *yA* with a of a *yA* allele possessing an intron scar from transient disruption (Figure 2.3). The *actA* and *hhtA* genes where used for normalization. One colony from each of the recovered *yA* disruptions (NID2014 and NID2015 with a loxP scar in second intron and NID2016 and NID2017, with a loxP scar in third intron) , were purify streaked, RNA was purified and used for qRT-PCR. Three out of the four, showed expression levels similar to the wildtype *yA* gene, while the fourth, NID2017, with a loxP scar in the third intron showed a somewhat lower expression levels. However, the NID2017 RNA was also of lower quality than the other samples. While the data does not conform to MIQE [12], this strongly suggests that this method of transient disruption is feasible.

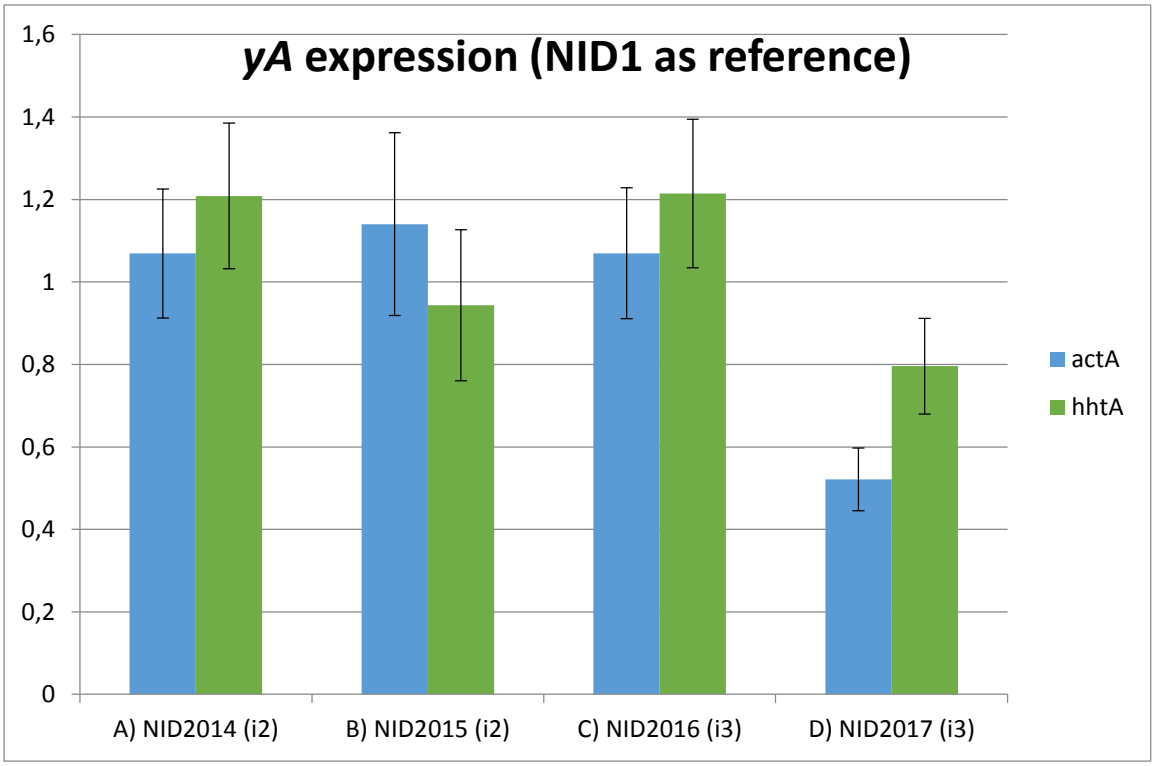

Figure 2.3 Test of *yA* expression in four strains in which *yA* have been restored. A) and B) shows the expression of *yA* in different transformant formerly disrupted in the *yA* second intron, while C) and D) are from two different transformants formerly disrupted in the third intron. Numbers are relative to *yA* expression in NID1 and two different household gene were used for normalization.

A Self-excising intron disruption cassette

While the above experiment provided proof of concept that introns could be used for transient disruptions of genes, it would be convenient if it did not require an additional cycle of protoplastation and transformation. Two DNA constructs, pyA-i2::tet-cre-pyrG and pyA-i3::tet-cre-*pyrG* were made containing, the cre recombinase under control of the fungal optimized tetON promoter [13] and a *pyrG* selection marker. These three elements were flanked by loxP sites in direct orientation and targeted into either the second or third intron, the exact same location as above. Linearized vector was transformated into NID1, disrupting either second or third intron and both transformations readily yielded yellow colonies (Figure 2.4, panel A and B), and when purify streaked they remain phenotypically stable, resulting in NID2018 and NID2019 respectively. However, when grown over long time, more than three weeks, few green spores could be seen in the microscope in the yellow colonies, which could potentially be due to either leakage of the promoter or a direct repeat recombination event in the *tetON* promoter, leaving only the *gpdA*

promoter causing constitutive expression. This is possible since the A. nidulans promoter is used both driving expression of the trans-activator and expression of the gene of choice [13], creating a direct repeat. When yellow colonies where transferred to non-selective media with doxycycline, inducing the *tetON* promoter, the colonies having second intron disrupted remained yellow, whereas those disrupted in the third intron of *yA* readily turned green (Figure 4, panel C and D). One colony of each were purify streaked and saved as NID2020 and NID2021.Considering that what will be left from the loxP-*pyrG*-loxP and the loxP-*tetON-cre-pyrG*-loxP, after cre mediated excision, is exactly the same, loxP, it was strange that one of the second intron disruption constructs readily worked while the other did not. Six colonies from the second intron disruption was transferred from the doxycycline containing media to MM-arg, however only a single one was able to grow, suggesting that *pyrG* had been removed from the others and that for the most part incomplete excision of the insert was not the reason for the inability to restore function of the disrupted gene.

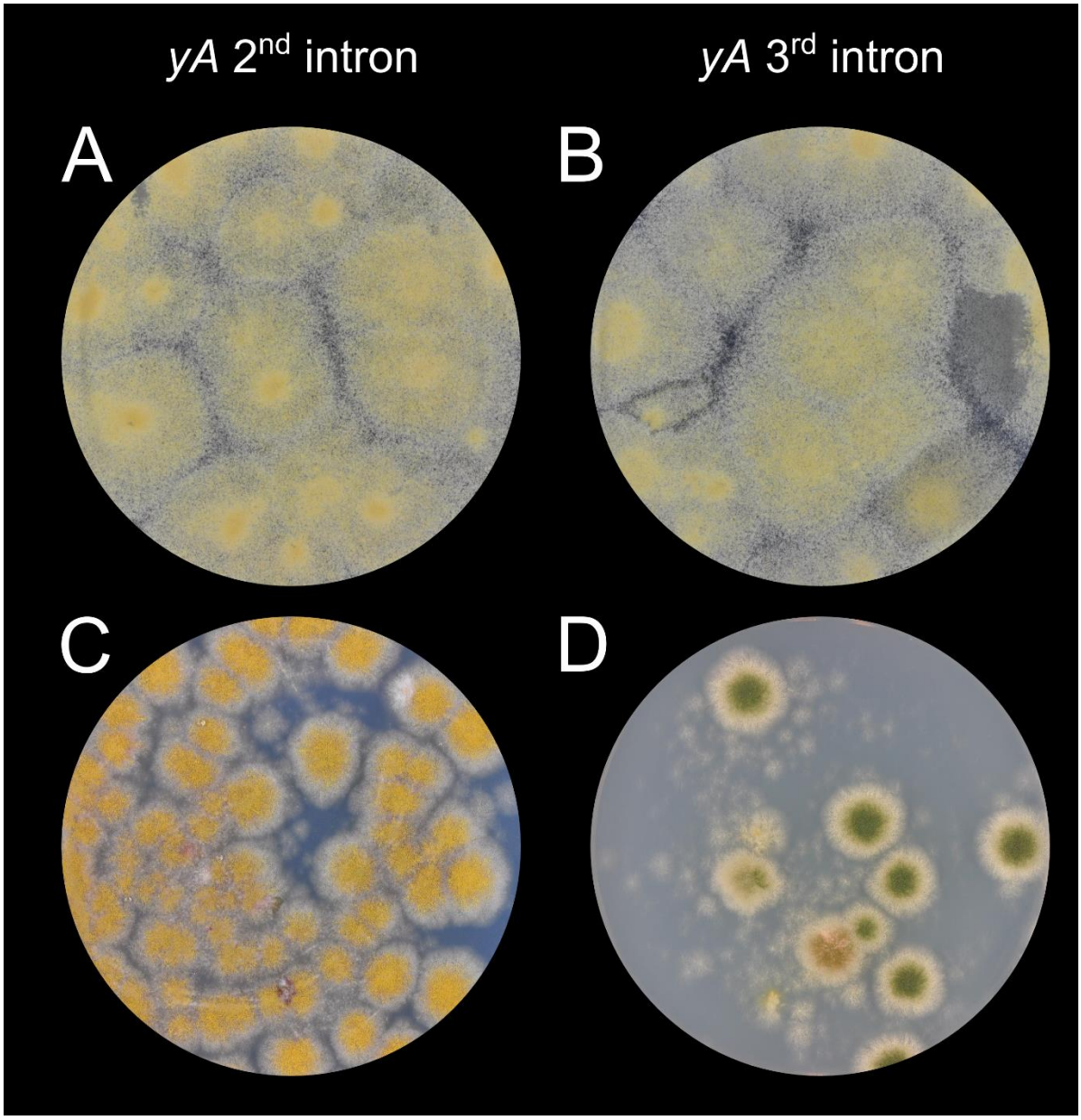

Figure 2.4 Self-excising intron disruption cassette. A) and B) shows transformations in which the second or third intron of *yA*, respectively, have been disrupted with the *cre* recombinase gene under control of the *tetON* promoter and *pyrG* selection marker. C) and D) shows colonies from A) and B) respectively, which after streak purifying are plated on non-selective media containing doxycycline, inducing the *tetON* promoter.

Transplanting introns to other genes

To figure out how far this concept can be pushed we decided to see if is possible to insert or replace one intron with another intron in the *A. nidulans wA* gene, which if disrupted gives rise to a white spored phenotype. We used the vector containing the third intron of *yA* with the loxP-tetON-cre-*pyrG*-loxP insert as template and constructed four vectors either inserting the intron at a location in

wA which does not contain an intron or tried to replace an existing intron with the modified *yA* intron. Two vectors allowing for insertion on the *yA* intron either between the 86th and the 87th bp or the 1933th and the 1934th of the *wA* gene were made, pwA-II1 and pwA-II2, while two other vectors, replacing either the first or the third intron of the same gene, pwA-IR1 and pwA-IR3 were similarly made.

All four vectors were transformed into NID1. Transformants from all four transformations displayed a white phenotype consistent with *wA* disruption (Figure 2.5, panel A-D), and they were purified streaked unto selective media, and one colony of each saved as NID2022, NID2023, NID2024 and NID2025. Afterwards each of the colonies and one other from the same transformation plate, were streaked on non-selective media (MM-arg-uri-ura) containing doxycycline, to see if excising out the tet-cre-pyrG could restore *wA* function, leading to a green phenotype. The strain carrying the construct for replacing the first intron in *wA* with the third intron of *yA* readily turned green again, while the other three strains remained white-spored (Figure 5, panel E-H). Colonies from all eight were purified streaked onto selective MM-arg media, to see whether failure to restore *wA* function could be result of a failure to excise the tet-cre-pyrG cassette, which could be indicated by an ability to grow on media not containing uridine and uracil. Two of each of the doxycycline exposed colonies were stabbed to minimal media containing arginine, but no uridine or uracil. The colonies which successfully had *wA* function restored, unsurprisingly did not grow since the *pyrG* marker has to be excised for intron splicing to be restored, however for the remaining colonies, one of each transformants grew, while the others did not.

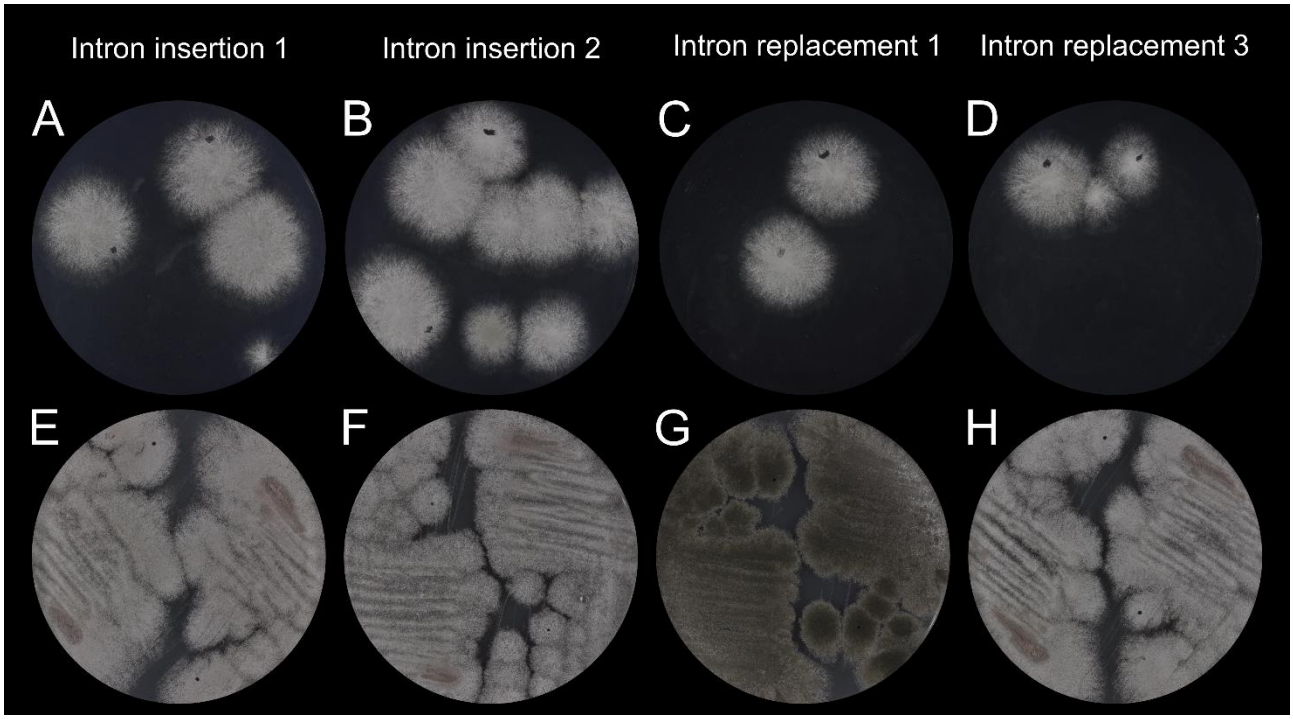

Figure 2.5 Intron transplantation. A) and B) shows transformation plates where NID1 is transformed with the vectors pwA-ii1 and pwA-ii2, inserting the third *yA* intron with the *tetON-cre-pyrG* insert into two different sites in wA. C) and D) shows transformation plates of NID1 transformed with the vector pwA-ir1 and pwA-ir3, where either the first or the third intron of *wA* is replaced with the third *yA* intron with the *tetON-cre-pyrG* insert. E), F), G) and H) shows colonies from A), B), C) and D) respectively, which after streak purifying are plated on non-selective media containing doxycycline.

Creating a strain transiently disrupted in pyrG and nkuA

One potential application of intron based transient disruption, would be to simultaneously disrupt *pyrG* and n*kuA*, the *ku70* homolog in *A. nidulans*, to get a strain with suitable traits for genetic engineering, with high frequencies of homologous recombination (HR) and a counter-selectable selection marker, but which could be cleaned up afterwards. While including the cre gene under control of the *tetON* promoter was an option, we decided instead to rely on later introduction of cre, to avoid potential issues with leakiness.

First a vector was made, pNid-pyrG::ble containing the bleomycin marker gene flanked by loxP sites and sequences for targeting the one intron in the *A. nidulans pyrG* gene. The construct was transformed into NID730, which is wildtype except for the *veA1* mutation[14], selecting for bleomycin resistance. Resulting colonies were transferred to 5-FOA containing media. Of 15 colonies transferred to 5-FOA, one had the ability to grow on the media and this transformant was

confirmed by PCR, and after streak purifying the resulting strain named NID1763. A challenge for testing whether the disruption is transient or not, is that *pyrG* is the only easily available selection marker. This makes it difficult to discern whether the *pyrG* prototrophy originates from the original *pyrG* gene being restored or from the *pyrG* gene introduced with the cre recombinase vector. The resulting strain, NID1763 was protoplasted and transformed with an AMA1 based vector, pCCM2 containing the cre recombinase encoding gene under control of the *A. nidulans gpdA* promoter and a *pyrG* selection marker. After transformation, selecting for uridine and uracil prototrophy, four transformants were two times re-streaked without selection to lose the AMA1 plasmid, since it is very unstable, even with selection it is readily being lost [15]. Afterwards the colonies were transferred to both minimal media and 5-FOA. All four colonies grew on MM, but not on 5-FOA. If the uridine/uracil prototrophy originated from the *pyrG* gene on the AMA1 plasmid it should have been lost during growth without selection, and even if still present, it should readily be lost upon growth on 5-FOA. Instead if the uridine/uracil prototrophy was due to the native *pyrG* gene being restored, it should be able to grow even after growth without selection, while when transferred to 5-FOA, without a way to get rid of the *pyrG* activity, it would be unable to grow. Furthermore, PCR confirmed that the bleomycin marker gene had indeed been excised. One colony was saved as NID2026. This strongly indicates that the *pyrG* gene could be transiently disrupted and then subsequently restored.

Next vectors, pNid-nkuA::hyg-pyrG-up and pNid-nkuA::pyrG-dw for transient disruption of *nkuA* was constructed. Since the strain was now *pyrG* deficient, *pyrG* could now be used for selection and afterwards recycled if flanked by direct repeats. However, since it is unclear how large an insert is needed to disrupt intron splicing, and the repeat left behind is only approximately 500 bp the hygromycin marker gene was additionally inserted, as filler DNA. To improve gene targeting frequencies, and since there is no obvious phenotype for *nkuA* disruption, rather than constructing one vector, two vectors were made for bipartite gene targeting [16]. One containing an upstream sequence for targeting the second intron of *nkuA*, the hygromycin marker, the first repeat and part of the *A. fumigatus pyrG* gene, and the other vector containing the other part of *pyrG* with some overlap, the second repeat and the downstream sequence for targeting *nkuA*. The strain with the transient *pyrG* disruption, NID1763, was transformed with both vectors, selecting for uridine/uracil prototrophy. In total 16 colonies were purify streaked and analyzed by PCR. A single colony showed the bands corresponding to a correct integration event. The transformant was plated on 5-FOA to excise the *pyrG* marker and a colony was isolated. It was checked by PCR that the new size

of the insert was consistent with the hygromycin marker gene and a single repeat from *pyrG*. The new strain, NID1764 was protoplasted and transformed with a vector for deleting *yA*, p6f-yA [17]. The transformation resulted in 12 yellow-spored colonies and a single green-spored one, indicating a gene targeting frequency above 90%, consistent with the phenotype also seen when deleting *nkuA*. The strain, NID1764 was likewise transformed with pCCM2 and the successful excision on the inserts in both nkuA and *pyrG* was confirmed by PCR. Unfortunately, due to time-constrains I did not get to test whether the *nkuA* disruption were actually reversible or not, and such whether the resulting strain, NID2027 was NHEJ proficient again.

Discussion

The results presented here demonstrate that the concept for transient disruption by insertion into an intron is certainly viable. One thing that remains to be seen is the effect when multiple disrupted genes have to be restored at once, however at least in the yeast *S. cerevisiae* multiple simultaneously excision events by cre-lox is possible [18]. As such there is good reason to believe it will also work in *Aspergillus*, even if it also gets expanded for instance to include more selection markers such as *argB*. A potential risk, if the genes are located on different chromosomes, is that multiple excision events can cause a cross-over event, and thus chromosomal rearrangements. However, since the disruption is from an intron and such in the middle of a gene, cross-over events would still result in the disrupted phenotype, and for instance in the case of *pyrG* such an event is easy to detect.

An important consideration is what the long term effects of such an inactivation are. Since the only loxP scar left is in the intron, the protein is completely unaffected, with no effects on either folding or localization. Preliminary qRT-PCR data on the restored *yA* gene, suggest that mRNA levels are not affected, and while analysis will have to be carried out for *pyrG* and *nkuA* and potential other targets, it is definitely promising.

One interesting aspect is that the length of the insert seems to be of importance. When disrupting the second intron of *yA*, the exact same location that tolerated the insert consisting of *pyrG* alone (1474 bp) for reversible disruption, did not tolerated the longer insertion consisting of the tetON-*cre-pyrG* construct (5072 bp), despite a majority of the colonies showing uridine and uracil auxotrophy. The auxotrophy suggests that the *pyrG* marker has been excised. Knowing why some

inserts can be tolerated and other not, would be very helpful. Sequencing of the site in some of those colonies that show the disruptant phenotype, but which still also shows uridine/uracil auxotrophy could shed light on why it does not work. The same is the case for the scenario where the replacement of one introns of *wA* with the third intron of *yA*, replacement of the first *wA* intron worked, while replacement of the third *wA* did not work.

An alternative use for this transient inactivation, is to use it for transient insertion. For instance, having Cas9 inserted in the genome for genome editing could be very useful. However for a production strain, having a DNA binding protein permanently inserted, might not be desirable. As such, transiently inserting Cas9 into an intron could be an elegant solution.

Ultimately while the concept still has some kinks it would be helpful to sort out, there is good reason to believe that this can be a valuable addition to the genetic engineering toolbox in *Aspergillus*.

Materials and methods

Strains and media

Escherichia coli strain DH5α was used to propagate all plasmids. The *Aspergillus nidulans* strains used in this study are listed in Table 2.1. Genomic DNA (gDNA) from fungal strains were isolated via FastDNA SPIN Kit for Soil DNA extraction kit (MP Biomedicals, USA). The mutant strains made in this study are also listed in Table 2.1. All strains were cultivated on standard solid glucose based minimal medium (MM) (1% glucose, 1x nitrate salt solution [19], 0.001% Thiamine, 1x trace metal solution [20], 2% agar), supplemented with 10mM uridine (Uri), 10mM uracil (Ura), and/or 4mM L-arginine (Arg) when required. Solid plates containing 5-fluoroorotic acid (5-FOA) were made as MM+Arg+Uri+Ura supplemented with filter-sterilized 5-FOA (Sigma-Aldrich) to a final concentration of 1.3 mg/ml. For transformation media (TM) glucose was replaced with 1M sucrose.

Table 2.1 List of strains used in the study

Strain name	Genotype	Source
NID1	*argB2, veA1, pyrG89, nkuAΔ*	EMCB strain collection[1]
NID730	*veA1*	EMCB strain collection[1]
NID1763	*veA1, pyrG::ble*	This study
NID1764	*veA1, pyrG::ble, nkuA::hph*	This study
NID2010	*argB2, veA1, pyrG89, nkuAΔ, yA::pyrG(i2)*	This study
NID2011	*argB2, veA1, pyrG89, nkuAΔ, yA::pyrG(i2)*	This study
NID2012	*argB2, veA1, pyrG89, nkuAΔ, yA::pyrG(i3)*	This study
NID2013	*argB2, veA1, pyrG89, nkuAΔ, yA::pyrG(i3)*	This study
NID2014	*argB2, veA1, pyrG89, nkuAΔ, yA$_{(i2\text{-}loxP)}$*	This study
NID2015	*argB2, veA1, pyrG89, nkuAΔ, yA$_{(i2\text{-}loxP)}$*	This study
NID2016	*argB2, veA1, pyrG89, nkuAΔ, yA $_{(i3\text{-}loxP)}$*	This study
NID2017	*argB2, veA1, pyrG89, nkuAΔ, yA$_{(i3\text{-}loxP)}$*	This study
NID2018	*argB2, veA1, pyrG89, nkuAΔ, yA::tetON-cre-pyrG(i2)*	This study
NID2019	*argB2, veA1, pyrG89, nkuAΔ, yA::tetON-cre-pyrG(i3)*	This study
NID2020	*argB2, veA1, pyrG89, nkuAΔ, yA$_{(i2\text{-}loxP)?}$*	This study
NID2021	*argB2, veA1, pyrG89, nkuAΔ, yA$_{(i3\text{-}loxP)}$*	This study
NID2022	*argB2, veA1, pyrG89, nkuAΔ, wA::tetON-cre-pyrG(ii1)*	This study

NID2023	*argB2, veA1, pyrG89, nkuAΔ, wA::tetON-cre-pyrG(ii2)*	This study
NID2024	*argB2, veA1, pyrG89, nkuAΔ, wA::tetON-cre-pyrG(ir1)*	This study
NID2025	*argB2, veA1, pyrG89, nkuAΔ, wA::tetON-cre-pyrG(ir3)*	This study
NID2026	*veA1, pyrG$_{(i1\text{-}loxP)}$*	This study
NID2027	*veA1, pyrG$_{(i1\text{-}loxP)}$, nkuA$_{(i2\text{-}loxP)}$*	This study

[1] Strain collection of the Eukaryotic Molecular Cell Biology group, DTU Systems Biology, Contact person: Uffe Hasbro Morten

PCR and USER cloning

All vectors were constructed by PCR and USER cloning as described by Nødvig et al [17]. Detailed vector construction can be found in Appendix 2.1 and the primer used in Appendix 2.2.

Placement of insert in introns

To be on the side of caution, it was ensured that the first seven and the last six bps of the intron were avoided (consensus donor and extended acceptor sites are six and five bp in length respectively [11]). Likewise, all sequences matching the motive NNCTRAY were avoided to protect the splice site. If no motive was found matching NNCTRAY, NNYTRAY was used instead. Since the potential splice site in general is located close to the acceptor site, in all cases the insert was placed between donor and splice site.

Transformation and strain validation by Tissue-PCR

Protoplastation were performed as described by Nielsen et al [16]. Transformation using either *pyrG* or *argB* as genetic marker was performed as described by Nødvig et al [17], while for transformation utilizing bleomycin selection, 10^7 protoplasts and ~3 µg of digested DNA were incubated for 15 min on ice, then 1 mL PCT was added and the mix incubated for 15 min at room temperature. 12 µg/ml bleocin B (InvivoGen, USA) was added to 15 ml molten 1M sorbitol based TM (TMsh; ~45°C, pH 6.5), and immediately poured into an empty 9 cm petri dish. After 24 h incubation at 30°C, an overlay of 15 ml TMsh was added. All candidate transformants were streak purified prior to verification. All strains which were verified by PCR, was done by tissue-PCR

analysis using mycelium as described by Nødvig et al [17]. The use of primers is described in Appendix 2.1 and the primer sequences can be found in Appendix 2.2.

Primers for gene disruption analysis were designed in two ways. One primer binds completely outside the sequence included in the gene targeting vectors, and the other binds either inside the insert or on the opposite side of the insert sequence. One setup detects only correct integration while the other detects both correct and failed integrations as well as heterokaryons.

qRT-PCR

RNA extraction and qRT-PCR was performed as described by Hansen et al [21]. RNA extraction was done by RNeasy plant kit (Qiagen, Netherlands) on five strains (Nid1 and two of each colonies where *yA* activity restored after disruption in second or third intron). 10 µg of RNA was treated with DNase I (Qiagen) and 1 µg was used for cDNA amplification by Phusion RT-PCR kit (Finnzymes, Finland). Afterwards quantitative reverse transcription-PCR (qRT- PCR) was performed in a Chromo 4 detector/PTC-200 apparatus (MJ Research, Canada) by using SYBR Green JumpStart Taq ReadyMix (Sigma, Germany). The *A. nidulans* actin gene, *actA AN6542* and the histone 3 gene, *hhtA AN0733* were the internal standards for normalization of expression levels.

Three primer pairs were used, Any/A-qRT-3-F/ ANyA-qRT-3-R, ANhhtA-RT-R/ ANhhtA-RT-F2 and ANactA-3'-F/ ANactA-3'-R, which can be found in Appendix 2.2. Samples were run in triplicates. The program was 94°C for 2 min and 40 cycles of 94°C for 10 s, 60°C for 15 s, and 72°C for 30 s. A melting curve from 65°C to 95°C with reads every 0.2 min ended the program to evaluate the purity of the reaction products. The fluorescence threshold values (CT) were determined by using the Opticon Monitor 3.1 software (MJ Research). The relative expression levels were approximated based on $2^{-\Delta\Delta C_T}$, with $\Delta\Delta C_T = \Delta C_{T(normalized)} - \Delta C_{T(calibrator)}$, where $\Delta C_{T(normalized)} = C_{T(target\ gene)} - \Delta C_{T(household\ gene)}$.The calibrator C_T values are those for the reference strain, NID1.

References

1. Nielsen ML, De Jongh WA, Meijer SL, Nielsen J, Mortensen UH. Transient marker system for iterative gene targeting of a prototrophic fungus. Appl Environ Microbiol. 2007;73: 7240–7245. doi:10.1128/AEM.01839-07

2. Nielsen JB, Nielsen ML, Mortensen UH. Transient disruption of non-homologous end-joining facilitates targeted genome manipulations in the filamentous fungus Aspergillus nidulans. Fungal Genet Biol. 2008;45: 165–170. doi:10.1016/j.fgb.2007.07.003

3. Sternberg N, Hamilton D. Bacteriophage P1 site-specific recombination. J Mol Biol. 1981;150: 467–486. doi:10.1016/0022-2836(81)90375-2

4. Schwartz CJE, Sadowski PD. FLP Recombinase of the 2 μm circle plasmid of Saccharomyces cerevisiae bends its DNA target. J Mol Biol. 1989;205: 647–658. doi:10.1016/0022-2836(89)90310-0

5. Schwartz CJE, Sadowski PD. FLP protein of 2 μ circle plasmid of yeast induces multiple bends in the FLP recognition target site. J Mol Biol. 1990;216: 289–298. doi:10.1016/S0022-2836(05)80320-1

6. Canosa I, Rojo F, Alonso JC. Site-specific recombination by the beta protein from the streptococcal plasmid pSM19035: minimal recombination sequences and crossing over site. Nucleic Acids Res. 1996;24: 2712–7. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=146011&tool=pmcentrez&rendertype=abstract

7. Sauer B. Functional expression of the cre-lox site-specific recombination system in the yeast Saccharomyces cerevisiae. Mol Cell Biol. 1987;7: 2087–2096. doi:10.1128/MCB.7.6.2087.Updated

8. Kühn R, Torres RM. Cre/loxP recombination system and gene targeting. Methods Mol Biol. 2002;180: 175–204. doi:10.1385/1-59259-178-7:175

9. Cheng TH, Chang CR, Joy P, Yablok S, Gartenberg MR. Controlling gene expression in yeast by inducible site-specific recombination. Nucleic Acids Res. 2000;28: E108.

10. Le Y, Sauer B. Conditional gene knockout using Cre recombinase. Mol Biotechnol. 2001;17: 269–275. doi:10.1385/MB:17:3:269

11. Kupfer DM, Drabenstot SD, Buchanan KL, Lai H, Zhu H, Dyer DW, et al. Introns and Splicing Elements of Five Diverse Fungi Introns and Splicing Elements of Five Diverse Fungi †. Eukaryot Cell. 2004; doi:10.1128/EC.3.5.1088-1100.2004.

12. Bustin S a, Benes V, Garson J a, Hellemans J, Huggett J, Kubista M, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem. 2009;55: 611–22. doi:10.1373/clinchem.2008.112797

13. Meyer V, Wanka F, van Gent J, Arentshorst M, van den Hondel C a MJJ, Ram AFJ. Fungal Gene Expression on Demand: an Inducible, Tunable, and Metabolism-Independent Expression System for Aspergillus niger. Appl Environ Microbiol. 2011;77: 2975–2983. doi:10.1128/AEM.02740-10

14. Mooney JL, Hassett DE, Yager LN. Genetic analysis of suppressors of the veA1 mutation in

Aspergillus nidulans. Genetics. 1990;126: 869–874. Available: http://www.ncbi.nlm.nih.gov/pubmed/2076818

15. Gems D, Johnstone IL, Clutterbuck AJ. An autonomously replicating plasmid transforms Aspergillus nidulans at high frequency. Gene. 1991;98: 61–67. doi:10.1016/0378-1119(91)90104-J

16. Nielsen ML, Albertsen L, Lettier G, Nielsen JB, Mortensen UH. Efficient PCR-based gene targeting with a recyclable marker for Aspergillus nidulans. Fungal Genet Biol. 2006;43: 54–64. doi:10.1016/j.fgb.2005.09.005

17. Nødvig CS, Nielsen JB, Kogle ME, Mortensen UH. A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi. Yu J-H, editor. PLoS One. 2015;10: e0133085. doi:10.1371/journal.pone.0133085

18. Jensen NB, Strucko T, Kildegaard KR, David F, Maury J, Mortensen UH, et al. EasyClone: method for iterative chromosomal integration of multiple genes in Saccharomyces cerevisiae. FEMS Yeast Res. 2014;14: 238–248. doi:10.1111/1567-1364.12118

19. Kaminskyj SGW. Fundamentals of growoth, storage, genetics and microscopy of Aspergillus nidulans. Fungal Genet Newsl. 2001; 25–31.

20. Cove DJ. The induction and repression of nitrate reductase in the fungus Aspergillus nidulans. Biochim Biophys Acta. Elsevier Publishing Company, Amsterdam; 1966;113: 51–56. doi:10.1016/S0926-6593(66)80120-0

21. Hansen BG, Salomonsen B, Nielsen MT, Nielsen JB, Hansen NB, Nielsen KF, et al. Versatile enzyme expression and characterization system for Aspergillus nidulans, with the Penicillium brevicompactum polyketide synthase gene from the mycophenolic acid gene cluster as a test case. Appl Environ Microbiol. 2011;77: 3044–3051. doi:10.1128/AEM.01768-10

Appendix 2.1 Vector construction

All vectors were assembled by USER cloning [1]. The vectors for insertion in the *yA* second intron were made in two steps. First part, containing the two targeting sequences with an AsiSI/Nb.BtsI between two loxP sites. Upstream targeting sequence was made by PCR amplification with CSN18+CSN19 and downstream by primers CSN48+CSN21 both from *A. nidulans* gDNA and both of these cloned into vector p1, assembling the loxP sites and the AsiSI/Nb.BtsI cassette in the USER primer tails. Similarly, a vector for insertion into the third intron of *yA* was made by amplification with primers CSN18+CSN49 for upstream sequence and primers CSN50+CSN21 for the downstream. These two vectors were then opened for further USER cloning by digestion with AsiSI and Nb.BtsI. For insertion of *pyrG* alone, *pyrG* from *A. fumigatus* was amplified from the vector pDEL2 [2] with the primers CSN22+CSN23 and cloned into both vectors. For the vectors also containing the cre recombinase and the tetON promoter, *pyrG* was amplified with primers CSN73+CSN23, the tetON promoter was amplified from pVG4-1 [3] in two parts with primers CSN59+96 and CSN97+CSN60, while the cre gene was amplified with primers CSN61+CSN62 from template APC161. All four fragments were cloned into both vectors.

For insertion of the third *yA* intron, with the tetON-cre-pyrG insert, into *wA*, the intron with insert was amplified from vector pyA-i3::tet-cre-pyrG, using primers CSN220+CSN96 and CSN97+CSN221. Gene targeting flanks for replacing either first or third intron of *wA* with the modified *yA* intron, was made by PCR using primers CSN276+CSN274 and CSN275+CSN277 for the first intron and primers CSN436+CSN434 and CSN435+CSN437 for the third intron, all amplified from *A. nidulans* gDNA. For the insertion of the modified *yA* intron as an extra intron in *wA*, flanks for the first insertion was made with primers CSN276+CSN432 and CSN433+CSN277, while the second insertion was made with primers CSN275+223 and CSN224+225 amplified from gDNA.

For creating a vector for disruption of *pyrG* with the bleomycin resistance gene, upstream targeting sequence was amplified from *A. nidulans* gDNA, with primers CSN110 and CSN177, while the downstream targeting sequence was amplified CSN178 and CSN113. The bleomycin resistance gene was amplified from pAN8-1 [4] with primers CSN171 and CSN172, and the loxP sites assembled by the USER tails.

To create the vectors for disruption of nkuA, 5' fragment and 3'of nkuA was amplified from gDNA with primers CSN51+CSN230 and CSN231+CSN54, the hygromycin marker gene was amplified from CSN171+CSN217 using pAN7-1 [5] as template, 5' and 3' parts of *pyrG* was amplified by CSN73+CSN298 and CSN297+CSN148 using pDEL2 [2] for template and repeats for direct repeat recombination was amplified with primers CSN215+CSN204 and CSN216+CSN205 using gDNA from S. cerevisiae CEN.PK113-11C as template. The 5' nkuA fragment, the hygromycin marker gene, one of the repeat and the 5' *pyrG* fragment was assembled in one vector and the 3' *pyrG* fragment, the other repeat and the 3' nkuA fragment in the other vector.

Mycelia PCR for strain verification

Correct insertion into the *pyrG* locus was verified using primers CSN162+CSN189 and primers CSN161+CSN189. Similarly, nkuA insertion was verified using primer CSN42+CSN190 and CSN155+CSN190

References

1. Nour-Eldin HH, Geu-Flores F, Halkier BA. USER cloning and USER fusion: the ideal cloning techniques for small and big laboratories. Methods Mol Biol. 2010;643: 185–200. doi:10.1007/978-1-60761-723-5

2. Nielsen JB, Nielsen ML, Mortensen UH. Transient disruption of non-homologous end-joining facilitates targeted genome manipulations in the filamentous fungus Aspergillus nidulans. Fungal Genet Biol. 2008;45: 165–170. doi:10.1016/j.fgb.2007.07.003

3. Meyer V, Wanka F, van Gent J, Arentshorst M, van den Hondel C a MJJ, Ram AFJ. Fungal Gene Expression on Demand: an Inducible, Tunable, and Metabolism-Independent Expression System for Aspergillus niger. Appl Environ Microbiol. 2011;77: 2975–2983. doi:10.1128/AEM.02740-10

4. Mattern IE, Punt PJ, van den Hondel C a MJJ. A vector of Aspergillus transformation conferring phleomycin resistance. Fungal Genet Newsl. 1988;35: 25–30.

5. Punt PJ, Oliver RP, Dingemanse MA, Pouwels PH, van den Hondel CAMJJ. Transformation of Aspergillus based on the hygromycin B resistance marker from Escherichia coli. Gene. 1987;56: 117–124. doi:10.1016/0378-1119(87)90164-8

Appendix 2.2 Primer table

Primers used in the study		
Primer name		**Sequence**
yA intron insertion		
CSN18	yA-i2-up-fwd	GGGTTTAAU ATGCTCGTCTTCCATCTCTGC
CSN19	yA-i2-up-loxP-AsiSI-Nb.BtsI-rv	ACTGCAT AACUTCGTATAGCATACATTATACGAAGTTATCGGAAAGGGCACACACTCTAA
CSN21	yA-i2-down-rv	GGTCTTAAUTCCCAAACATCAACCCCGT
CSN22	Af-pyrG-fwd	AGAGCGAUCGTGGAGTTACCAGTGATTG
CSN23	Af-pyrG-rv	TCTGCGAUCTTGCTAGATGACTGGTAGG
CSN48	yA-i2-down-fwd2	AGTTATGCAGUGAGAGCGATCGCAGACACTGC ATAACTTCGTATAATGTATGCTATACGAAGTTATCGTCCCATTTCTGCATGGACCTAA
CSN49	yA-i3-up-rv	ATCGCTCUCACTGCATAACTTCGTATAGCATACATTATACGAAGTTAT CTTCATCGTGCAGATAAGGTGTAAT
CSN50	yA-i3-down-fwd	AGAGCGAUCGCAGACACTGCATAACTTCGTATAATGTATGCTATACGAAGTTA TCCCATTGAGGCAGTCTTATTGAT
CSN59	tetON fwd	AGAGCGAUAGCTTCGGAGAATATGGAGCTTCATCGA
CSN60	tetON-rv	AAACGGUGATGTCTGCTCAAGC
CSN61	cre-fwd	ACCGTTU ATGTCCAATTTACTGACCGTACACC
CSN62	cre-rv	acgaagtcu AACCTTACCCAAGAGTTCGCC
CSN73	pyrG-U1-fwd	agacttcguCGTGGAGTTACCAGTGATTG
CSN96	tetON-int-rv	ATGAAGTTAAUGCATGGCAGACACTGAAGC
CSN97	tetOn-int-fwd	ATTAACTTCAUCCTACTGTCCTACCCGCAG
wA intron insertion or replacement.		

CSN220	yA-i3-fwd	AGTACGUCTGATCAGATTACACCTTATCT
CSN221	yA-3-rv	ACTATUGAGACATGATCAATAAGACTGC
CSN222	wA-up-fwd	GGGTTTAAU GGACCCATACCGTGTCTATCTCTT
CSN223	wA-up-rv	ACGTAC U ATCCATCGGCTCCGTCATC
CSN224	wA-dw-fwd	AATAG U TGCAGAGCAGATGGTGTTGG
CSN225	wA-dw-rv	GGTCTTAAU ACGGCTCTCAACGACTCTCTG
CSN274	Anid-wA-i3-ins-up-rv	ACGTAC U TGATAAAACATCCCAATTGATATATGC
CSN275	Anid-wA-i3-ins-dw-fwd	AATAG U TACTATGGTGACCTTGGTCATACATTC
CSN276	Anid-wA-up-fwd	GGGTTTAAU TGACGCCAGCTCTCTTCCC
CSN277	Anid-wA-dw-rv	GGTCTTAAU TCAGGAAACCGACCCGACA
CSN432	wA-II1-up-rv	ACGTAC UTCTTTGCCTGTAGCAGACGACG
CSN433	wA-II1-dw-fwd	AATAG UCACTCTCTCCTTTCGAGCTTTCTC
CSN434	wA-IRI3-up-rv	ACGTAC UGAAGGACAGAGTGCAGCTCCATT
CSN435	wA-IRI3-dw-fwd	AATAG UCTCTTTATGCAGATAGTGCTCAAACC
CSN436	wA-IRI3-up-fwd	GGGTTTAAU GGCATCATTAGGAATAACGCCAT
CSN437	wA-IRI3-dw-rv	GGTCTTAAU GCAAACACCAGATCGTCAGACAT
Transient *pyrG* and *nkuA* disruption		
CSN110	Nid-pyrG-i1-up-fwd	GGGTTTAAU CCCAGAAAATTGCCACATTAGACATTG
CSN177	Nid-pyrG-loxp-up-rv	ATACATTAUACGAAGTTAT TGGTGCAGGATGTACGGTCAGC

CSN178	Nid-pyrG-loxP-dw-fwd	ATGTATGCUATACGAAGTTAT ATGCCCCTCCAGGATAACAAATAGC
CSN113	Nid-pyrG-i1-dw-rv	GGTCTTAAUTTACTGGCAGGGTACACGGCTGA
CSN171	gpdAp-loxP-fwd	ATAATGTAUGCTATACGAAGTTAT GCGTAAGCTCCCTAATTGGC
CSN172	trpCt-loxP-rv	AGCATACAUTATACGAAGTTAT TTACCTCTAAACAAGTGTACCTGTGC
CSN51	nkuA-i2-up-fwd	GGGTTTAAUGGCACAGAGTTCGGCGTTG
CSN230	nid-nkuA-i2-up-rv	ATACATTAU ACGAAGTTATCGAACGGTCTGACTTTGATTCA
CSN231	nid-nkuA-i2-down-fwd	ATGTATGCU ATACGAAGTTATCGACCCTACTAATCACCAAATAGCACT
CSN54	nkuA-i2-down-rv:	GGTCTTAAUCCACCTCCCGTCCACAATC
CSN215	R1-L3-fwd	aagtgtaau CAAACATCTACACAATTAGCAAGGG
CSN216	R2-loxP2-rv	AGCATACAUTATACGAAGTTAT GTTACACGGAAGGAGAGCAGTAAG
CSN217	TtrpC-L3-rv	attacactu TTACCTCTAAACAAGTGTACCTGTGC
CSN204	R1-(L1)-rv	acgaagtcu GTTACACGGAAGGAGAGCAGTAAG
CSN205	R2-(L2)-fwd	aagtctacu CAAACATCTACACAATTAGCAAGGG
CSN73	pyrG-U1-fwd	agacttcguCGTGGAGTTACCAGTGATTG
CSN148	pyrG-rv	agtagactu CTTGCTAGATGACTGGTAGG
CSN297	Af-pyrG-int-PacI-fwd	GGGTTTAAU TGATGATACAGGTCTCGGTCCC
CSN298	Af-pyrG-int-PacI-rv	GGTCTTAAU GGAAGAGAGGTTCACACCCGT
qRT-PCR primer		
	ANyA-qRT-3-F	CCACATCCGATCCATAAGCAC

	ANyA-qRT-3-R	CATGCCCCATCCATGAGC
	ANhhtA-RT-R	GAGGCGACGAGCAAGCTG
	ANhhtA-RT-F2	GTGCTCTCCAGGAGTCCG
	ANactA-3'-F	GACGTCCGTAAGGATCTGTACG
	ANactA-3'-R	GCGGTGGACGATCGAAGG
Primers for PCR verification		
CSN155	nkuA-int2-fwd	AAGAGTCGGCGAGAAGTTGTTA
CSN161	Nid-pyrG-int1-fwd	ACAGCCCAAAGCCTACAAAT
CSN162	ble-int-fwd	CGAGATCGGCGAGCAGC
CSN189	Nid-pyrG out-rv	CAATCCCTTAAAGCCTTCCGT
CSN190	Nid-nkuA-out-rv	GAAGAAGAATCAGTGGATTGGGTC
CSN42	Af-pyrG-sq2-fwd	TGTAGTGCCAGTACGAGTGTTGTG

Chapter 3: A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi

PLOS | ONE

RESEARCH ARTICLE

A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi

Christina S. Nødvig, Jakob B. Nielsen, Martin E. Kogle, Uffe H. Mortensen*

Eukaryotic Molecular Cell Biology, Section for Eukaryotic Biotechnology, Department of Systems Biology, Technical University of Denmark, Søltofts Plads, Kongens Lyngby, Denmark

* um@bio.dtu.dk

Abstract

The number of fully sequenced fungal genomes is rapidly increasing. Since genetic tools are poorly developed for most filamentous fungi, it is currently difficult to employ genetic engineering for understanding the biology of these fungi and to fully exploit them industrially. For that reason there is a demand for developing versatile methods that can be used to genetically manipulate non-model filamentous fungi. To facilitate this, we have developed a CRISPR-Cas9 based system adapted for use in filamentous fungi. The system is simple and versatile, as RNA guided mutagenesis can be achieved by transforming a target fungus with a single plasmid. The system currently contains four CRISPR-Cas9 vectors, which are equipped with commonly used fungal markers allowing for selection in a broad range of fungi. Moreover, we have developed a script that allows identification of protospacers that target gene homologs in multiple species to facilitate introduction of common mutations in different filamentous fungi. With these tools we have performed RNA-guided mutagenesis in six species of which one has not previously been genetically engineered. Moreover, for a wild-type *Aspergillus aculeatus* strain, we have used our CRISPR Cas9 system to generate a strain that contains an *AACU_pyrG* marker and demonstrated that the resulting strain can be used for iterative gene targeting.

Citation: Nødvig CS, Nielsen JB, Kogle ME, Mortensen UH (2015) A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi. PLoS ONE 10(7): e0133085. doi:10.1371/journal.pone.0133085

Editor: Jae-Hyuk Yu, The University of Wisconsin - Madison, UNITED STATES

Received: March 17, 2015

Accepted: June 22, 2015

Published: July 15, 2015

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by Grant 11-116803 from The Danish Council for Strategic Research.

Competing Interests: The authors have declared that no competing interests exist.

Introduction

Filamentous fungi are ubiquitous organisms that impact human life in both positive and negative manners. For example, fungi play a major role in recirculating biomass in ecosystems, as they degrade basically all types of organic matter. For this reason they serve as a major source of industrially relevant enzymes, e.g. amylases, cellulases, lipases, pectinases, and proteases [1]. On the other hand, the same properties allow fungi to infect and deteriorate buildings, food and fodder and even living organisms, including humans, where they may cause fatal disease. Filamentous fungi also display elaborate secondary metabolisms that they use for chemical warfare, for signaling and as pigments. Many of these compounds are mycotoxins, which harm, or even kill, livestock and humans [2]. For example, several aspergilli produce carcinogenic aflatoxin and costs due to controlling mycotoxins in farm products amounts to several billion US$ per year [2,3]. However, fungal metabolites also include medically relevant

compounds like antibiotic penicillins, cholesterol lowering statins and immunosuppressive mycophenolic acid [3]. Importantly, the vast majority of fungal secondary metabolites await discovery and/or medical characterization, and the pool of fungal secondary metabolites is considered a large, and still under-exploited reservoir of drug leads and potential beneficial food additives.

Despite the impact of filamentous fungi on human life, detailed knowledge into the molecular biology and biochemistry is only available for a few model fungi and there is therefore much to explore. For that reason the number of fully genome sequenced fungal species is rapidly increasing, e.g. via projects like the 1000 Fungal Genomes Project or the Aspergillus Whole-Genus Sequencing project (http://genome.jgi.doe.gov/). The fact that genetic tools are scarce or non-existent for the vast majority of fungi greatly hampers the exploitation of these genome sequences to gain experimental insights into the biology of these fungi. Firstly, lack of genetic markers makes gene engineering difficult, and secondly, if functional markers are available, low gene-targeting frequencies makes reverse genetics tedious. In model fungi, mutations disabling *pyrG*, encoding orotidine-5'-phosphate decarboxylase, offer a classic strategy to establish a robust, selectable and counter-selectable selection marker that allows for multiple rounds of gene targeting [4,5]. However, introduction of such mutations in fungi for which no genetic tools are available can be laborious and time consuming and therefore constitute a major bottleneck in the genetic characterization of fully sequenced fungi.

The bacterial and archaeal immune mechanism CRISPR (clustered regularly interspaced short palindromic repeats)-Cas9 has recently been engineered into a powerful gene editing system [6–10]. Hence, a CRISPR-Cas9 system consisting of only two components, the Cas9 nuclease and a single chimeric guide RNA (sgRNA), allows for the introduction of specific DNA double strand breaks (DSBs), which in turn can be employed to efficiently stimulate gene targeting [11]. Accurate targeting of the RNA-guided Cas9 nuclease to a specific DNA sequence is achieved by the protospacer sequence of the sgRNA. The protospacer consists of only 20 nucleotides that recognize the target site by base pairing (Fig 1A). The fact that a synthetic oligonucleotide easily accommodates sequence corresponding to a protospacer makes it simple to construct and express genes encoding novel sgRNAs for Cas9. Programming of Cas9 to recognize new targets is therefore much easier than other systems that also function by inducing specific DSBs, such as TALENs or zinc-finger nucleases [12–14]. The only restriction of the CRISPR-Cas9 system is the requirement for a short Protospacer Adjacent Motif, PAM, adjacent to the binding site of the guiding protospacer sequence at the target site. The PAM sequence varies between Cas9 proteins [15], but for the *Streptococcus pyogenes* Cas9, which is commonly used for CRISPR-Cas9 gene editing, the very frequently occurring NGG sequences are accepted as PAMs (NAG is also accepted, but with reduced efficiency [16]. Hence, virtually all genes can be targeted by RNA-guided *Streptococcus pyogenes* Cas9. Like for TALENs and for zinc-finger nucleases, off-target effects have been reported for RNA-guided Cas9 gene editing, but a few mismatches scattered in the pairing region, especially in the 12 PAM proximal nucleotides of the protospacer, appears to eliminate mutagenesis [16,17]. Hence, off-target effects may mostly be problematic with organisms possessing large genomes. Indeed, whole genome sequencing of *Saccharomyces cerevisiae* strains that has been mutated by CRISPR-Cas9 indicated that off-target effects are insignificant in this species [18,19].

In higher eukaryotes, such as mammalian cell system or plants, where gene targeting is notoriously difficult, the CRISPR-Cas9 system has already had a huge impact and CRISPR-Cas9 has been successfully adapted to a wide range of organisms, such as yeast [20], mammalian cells [6,8], fish [7], and plants [21]. Here we present a CRISPR-Cas9 system adapted for filamentous fungi, which we have used to introduce specific mutations in six different species in two different sections of *Aspergillus*. For example, we have shown that it is possible to start

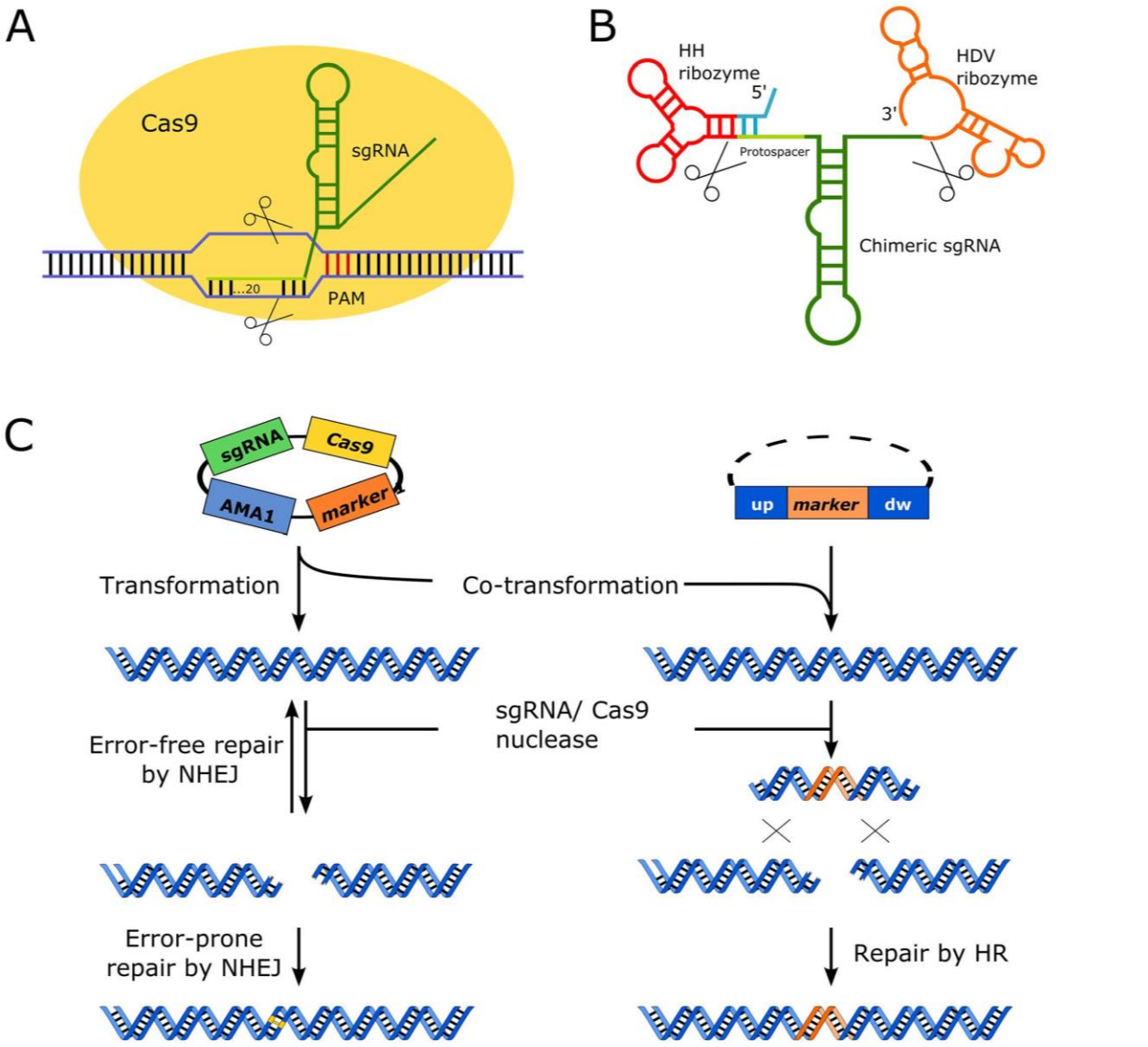

Fig 1. Components of the CRISPR-Cas9 system. Individual components in both panels are not drawn to scale. **A)** RNA guided nuclease Cas9 specifically cleaves a genomic target sequence. The protospacer and scaffold sections of the sgRNA are represented in light green and green, respectively. The protospacer basepairs with the complementary strand of the target sequence to form a D-loop. Note that efficient cleavage depends on the presence of the three bp PAM sequence in the target sequence, which is located directly downstream of the region invaded by the protospacer. Blunt-end cleavage occurs between bases located 3–5 bp upstream of the PAM, as indicated by scissors **B)** Liberation of the sgRNA, represented in light green/green, from a polymerase II transcript by intrinsic hammerhead (HH) ribozyme and hepatitis delta virus ribozymes (HDV) ribozyme represented in red/blue and orange respectively. Cleavage points are indicated by scissors. The blue part of HH basepairs with the protospacer for efficient cleavage. **C)** A fungal AMA1 based vector harboring Cas9 and sgRNA encoding genes are transformed into a fungus. The Cas9/sgRNA riboprotein induces a site specific DNA DSB. In the absence of a homologous template for HR repair (left lane), the DNA DSB is repaired by NHEJ; and error/prone repair by NHEJ results in mutations as indicated by yellow base pairs. In the presence of a linear or circular gene targeting substrate (right lane), specific Cas9 induced DNA DSBs may be repaired by HR resulting in a gene targeting event illustrated as insertion of an orange marker gene in this example.

doi:10.1371/journal.pone.0133085.g001

with a wild-type *Aspergillus aculeatus* strain and quickly introduce an *AACU_pyrG* mutation to allow for iterative gene targeting. We envision that these tools can be used to rapidly expand the repertoire of fungi where genetic engineering is possible and therefore contribute to accelerate the exploration and industrial exploitation of fungal biology.

Results

A versatile CRISPR-Cas9 system for genetic engineering of filamentous fungi

Our system for CRISPR-Cas9 mediated mutagenesis is designed for simple and versatile use in a broad spectrum of fungi. Hence, it allows directed mutagenesis by transforming a target host with a single plasmid that contains the genes encoding the two components of the CRISPR--Cas9 system, the Cas9 nuclease and a single chimeric guide RNA, see F 1. RNA guided DNA DSBs generated by Cas9 may be repaired in an error-prone manner by non-homologous end-joining (NHEJ) and directly lead to mutation; or alternatively, if a gene-targeting substrate is co-transformed into the strain together with the CRISPR-Cas9 vector, the DSBs may serve to increase the gene targeting efficiency at their site (See Fig 1C). To facilitate RNA guided mutagenesis in a broad range of filamentous fungi, four different CRISPR-Cas9 plasmids (see Fig 2) have been constructed each containing a commonly used fungal marker, *AFUM_pyrG*, *AN_argB*, ble^R or hyg^R, as well as the AMA1 sequence, which has been shown to support replication in many different fungal species [22,23].

Both components of the CRISPR-Cas9 system were modified to enhance functionality in fungi. Firstly, the *cas9* gene from *S. pyogenes* was codon optimized for expression in industrially relevant *A. niger* and 3'-extended by a sequence encoding a SV40 nuclear localization signal (PKKKRKV). The resulting synthetic gene was fused to the strong constitutive *A. nidulans tef1* promoter and to the *tef1* terminator[24,25]. The resulting fusion gene includes a consensus Kozak sequence [26] in front of the start codon. Secondly, the short sgRNAs do not contain a cap structure and poly A-tail, and in CRISPR-Cas9 systems they are therefore typically transcribed from genes controlled by RNA polymerase III promoters [27]. However, as these promoters are ill defined in filamentous fungi, we employed a variant of the CRISPR-Cas9 system in which the sgRNA is embedded in the middle of a larger transcript synthesized by RNA polymerase II [27]; the sgRNA is liberated from the larger transcript in the nucleus by the action of two ribozyme sequences, 5'-end hammerhead (HH) and 3'-end hepatitis delta virus (HDV), which flank the sgRNA, (Fig 1B). Robust synthesis of the transcript containing the sgRNA sequence is ensured by employing the strong constitutive *A. nidulans gpdA* promoter (*PgpdA*) and the *trpC* terminator (*TtrpC*). A similar strategy has previously been employed in other organisms, e.g. yeasts [27,28] and mammalian cells [29].

Construction of CRISPR-Cas9 vectors for directed mutagenesis of filamentous fungi

CRISPR-Cas9 vectors (see Fig 2 for details) with specific sgRNA genes are generated in a single USER fusion cloning step mediated by *E. coli*. The CRISPR-Cas9 vectors therefore contain an *E. coli* origin of replication, a B-lactamase gene for selection and a PacI/Nt.BbvCI USER cassette to facilitate insertion and construction of the sgRNA gene. Novel sgRNA genes controlled by the *PgdpA* and *TtrpC* sequences are inserted into the USER cassette by combining two PCR fragments amplified from plasmid pFC334 by USER fusion. As a consequence of the ribozyme based release strategy, there are two interdependent variable regions in the sgRNA gene. One region contains the 20 base pairs (bp) of the protospacer, the other region is a 6 bp inverted

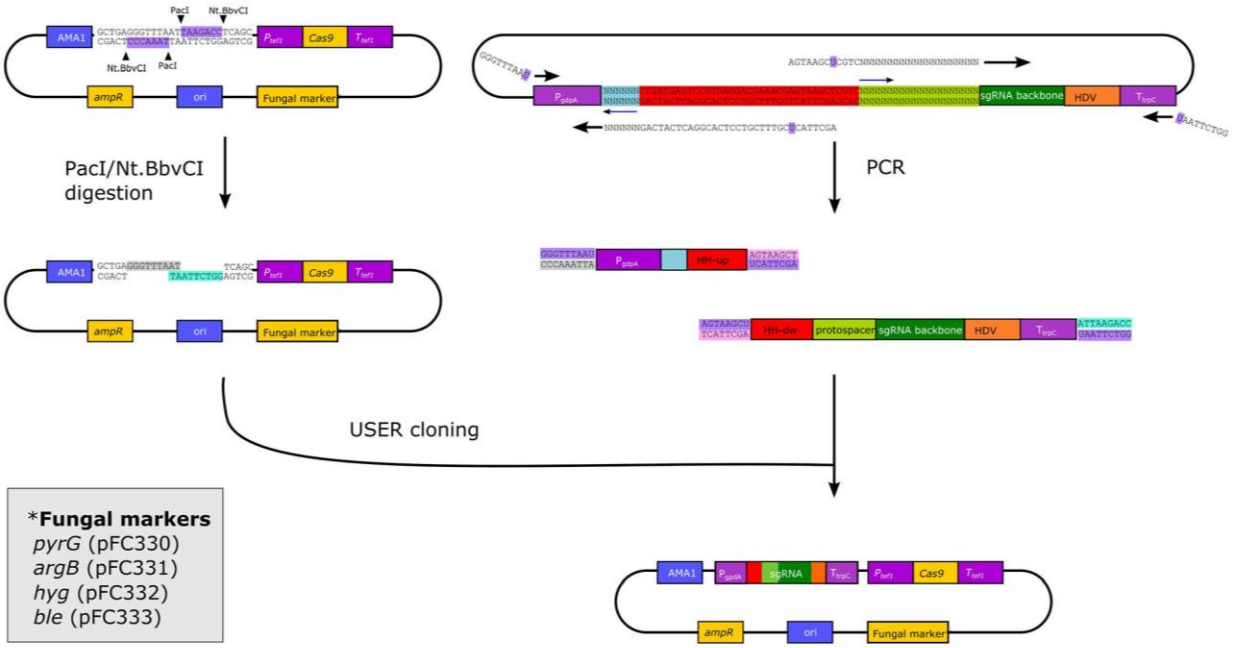

Fig 2. Construction of new CRISPR-Cas9 vectors for directed mutagenesis of filamentous fungi. Construction of fungal CRISPR-Cas9 vectors with variable sgRNA genes controlled by *gpdA* promoter and *trpC* terminator (no DNA elements are drawn to scale). The vector backbone for construction of new Fungal vectors for Cas9 induced genetic engineering are derived from the plasmid series pFC330-333. Sticky ends for USER cloning are achieved by opening the PacI/Nt.BbvCI USER cassette of pFC330-333 by the concerted action of restriction enzymes PacI and Nt.BbvCI (left side of panel). The two PCR fragments necessary for construction of the sgRNA gene, are both obtained by using pFC334 as template (right side of panel). This vector contains a protospacer for targeting yA (in light green), which is represented by 20 Ns indicating that it is not intended to match the primer; and in principle could be any sequence. The sections of the sgRNA gene encoding the variable parts of the transcript, the 20 bases of the protospacer (in light green) and the reverse complementing 6 bases of HH (in light blue), are introduced via tails added to the ends of the two primers that define the down- and upstream ends of the two PCR fragments, respectively. The position of the resulting inverted repeat located in the variable regions is indicated by blue arrows (top of panel). After amplification, the two PCR fragments are fused and inserted into vector pFC330-333 (Four variants exist) by USER cloning in a single step. For this purpose, each PCR fragment is generated by primers containing a tail with a uracil base (in purple). Elimination of the uracil bases in the PCR fragments by Uracil DNA glycosylase and DNA glycosylase-lyase Endonuclease VIII (USER Enzyme) results in the production of pairwise complementary overhangs at the ends of all fragments allowing selected ends to be fused in a directional manner. For simplicity, all complementary ends are visualized in the same color.

doi:10.1371/journal.pone.0133085.g002

repeat of the 5'-end of the protospacer, which needs to be present in the HH sequence, where it serves to complete the HH cleavage site, (Fig 1B). USER fusion allows for incorporation of the two variable regions into the sgRNA gene in a single cloning step. Specifically, this is achieved by extending the downstream primer used to generate the fragment containing the *gpdA* promoter with a 39 nucleotide long tail, and the upstream primer used to generate the sgRNA coding PCR fragment with a 32 nucleotide long tail. These two primers and the resulting PCR fragments are the only components that need to be acquired for the construction of new target-specific CRISPR-Cas9 vectors.

CRISPR-Cas9 efficiently introduces directed mutations into the *yA* gene of *A. nidulans*

To test for the functionality of our system, we first attempted to mutate the *yA* gene of *A. nidulans*. Successful mutagenesis of this gene is easy to monitor, as inactive *yA* alleles result in a yellow color of conidia (asexual spores) rather than the green color characteristic for wild-type conidia spores [30,31]. CRISPR-Cas9 plasmids containing either no sgRNA gene, or an sgRNA

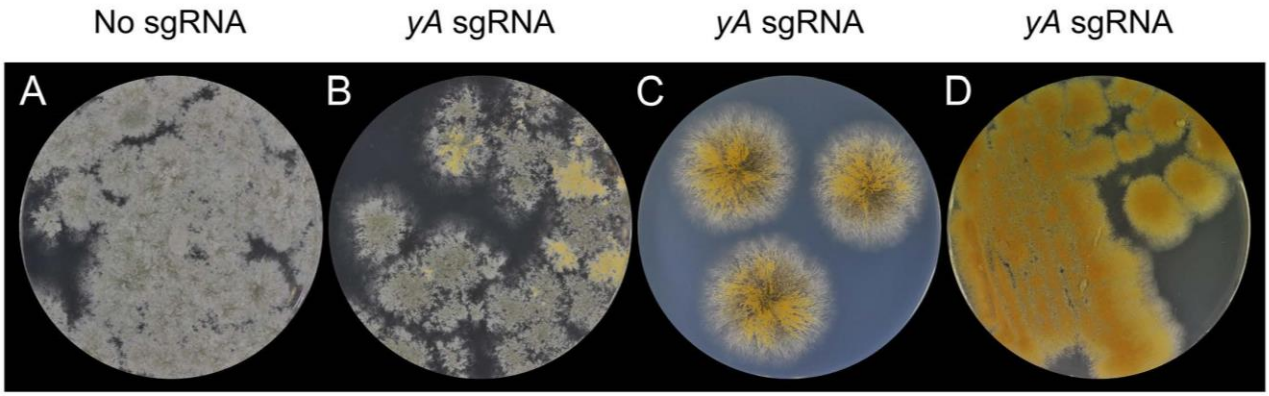

Fig 3. RNA guided Cas9 efficiently introduces directed mutations into the *yA* gene of *A. nidulans*. *A. nidulans* transformed with **A)** pFC331 encoding Cas9, but not a sgRNA, and with **B)** pFC334 encoding Cas9 and a *yA* specific sgRNA. **C)** Stabs of green transformants on solid selective medium from the plate shown in panel B. D) Stabs of yellow transformants on solid non-selective media. Despite loss of pFC334 they remain phenotypically stable

doi:10.1371/journal.pone.0133085.g003

gene with a protospacer targeting exon 4 of *yA* on chromosome I, were introduced into *A. nidulans* strain using *argB* as selection marker in three individual experiments. In the control experiments, where transformants expressed only *cas9*, the emerging colonies were all solid green (Fig 3A). In contrast, yellow colonies, or colonies with yellow sectors, were observed amongst transformants expressing both the *yA* specific sgRNA gene and *cas9* (Fig 3B) suggesting that the RNA-guided Cas9 nuclease indeed induces mutations in *yA*. As most of the colonies (approximately 70–80%) remained green on the primary transformation plates, we inoculated 12 green transformants onto solid selective medium to examine this phenotype in more detail. In all cases, each inoculation developed into yellow/green sectoring colonies with the vast majority of the colony area being yellow (Fig 3C and data not shown). This experiment shows that RNA-guided mutagenesis of *yA* is efficient in a growth dependent manner. Restreaking yellow colonies on medium without selection resulted in solid plasmid-free yellow colonies showing that the yellow phenotype is genetically stable in the absence of the *yA* specific sgRNA-Cas9 nuclease (Fig 3D). Next, to confirm that the yellow phenotype results from mutations in *yA* that are induced by the specific *yA*-sgRNA-Cas9 nuclease; we streak purified ten yellow colonies and generated PCR fragments covering the region of *yA* that was targeted by the nuclease. The ten PCR fragments were sequenced and in all ten cases mutations were identified at the site expected to be cleaved by Cas9. These results show that the mutations were likely obtained by error-prone NHEJ repair of the Cas9 induced break (Table 1). In eight out of the ten cases, small, one or two bp deletions, at the break created frameshifts in the *yA* gene explaining the lack of *yA* activity. Similarly, small indels have been reported as a frequent output of CRISPR-Cas9 mutagenesis in other species, e.g. in human HEK-293T cells [17]. The two remaining *yA* mutations contained 60 bps and 84 bps large inserts. Interestingly, these inserts are identical to two different loci on *A. nidulans* chromosome V. No sequences matching the protospacer-PAM sequence were identified in the regions 1 kb up- and downstream of the positions of these sequences. Moreover, the sequence with the highest identity to PAM and the adjacent 12 bps of the protospacer, which are most important for specificity, contained 7 mismatches, including one in the PAM sequence. These observations indicate that the 60 bps and 84 bps sequences were not captured due to additional RNA-guided Cas9 activity at these loci.

Table 1. Mutation spectrum of RNA guided Cas9 mutagenesis.

Species	Target Gene	Colony #	Mutation	Wild-Type Target Sequence[1]	Mutated Target Sequence [2]
A. nidulans	*yA*	1, 3, 4, 8, 9, 10	1 bp deletion	GGCGGAGTATCATAACATCG	GGCGGAGTATCATAAC-TCG
	yA	5, 6	2 bp deletion	GGCGGAGTATCATAACATCG	GGCGGAGTATCATAA—TCG
	yA	7	1 bp deletion, 60 bp insertion	GGCGGAGTATCATAACATCG	GGCGGAGTATCATAAC-60 bp[4]-TCG
	yA	2	84bp insert	GGCGGAGTATCATAACATCG	GGCGGAGTATCATAACA-84 bp[3]-TCG
A. aculeatus	*albA*	1	1 bp deletion	CGGTTCTTCAACATGTCGCC	CG-TTCTTCAACATGTCGCC
	albA	2	10 bp deletion	CGGTTCTTCAACATGTCGCC	—-TTCTTCAACATGTCGCC
	pyrG	1	1 bp deletion	CCCACATCATCAACTGCAGCATC	ACA-CATCAACTGCAGCATC
	pyrG	2	2 bp deletion	CCCACATCATCAACTGCAGCATC	ACA—ATCAACTGCAGCATC
A. niger	*albA*	1	83 bp deletion	AGTGGGATCTCAAGAACTAC	50—Protospacer—13
	albA	2	83 bp deletion	AGTGGGATCTCAAGAACTAC	50—Protospacer—13
A.. carbonarius	*albA*	1	7 bp deletion	AGTGGGATCTCAAGAACTACTGG	AGTGGGATCT———-TACTGG
	albA	2	24 bp deletion	AGTGGGATCTCAAGAACTACTGG	AGTGGGATCT-24bp
A. luchuensis	*albA*	1	70 bp deletion	AGTGGGATCTCAAGAACTACTGG	AGTGGGATCTCAAGAAC—70-
A. brasiliensis	*albA*	1	11 bp deletion	AGTGGGATCTCAAGAACTACTGG	AGTGGGATCT————-GG
	albA	2	25 bp deletion, 1 bp insertion	AGTGGGATCTCAAGAACTACTGGATCCCCTAT	AGTG——C ————————-AT

[1]Underlined bp shows expected location for Cas9 induced DSBs.

[2]Hyphens indicate deleted bp

[3]GCCATTGTTGGCTCGTGAAGTGTACGGATTGATGTATCGTCGTATCTGCATATTGCCCCTGAGACTGATGATCATGTCTGTCGGA; match AN11611

[4]TTCAAAATCTCGGAGGCTGATTGTTCCACGATGCGGGTGACGGCTCCTCGGGGCGTTTCT; match AN10634

doi:10.1371/journal.pone.0133085.t001

albA and *pyrG* in *A. aculeatus* can be efficiently mutated by CRISPR-Cas9

Next, as a first test of the versatility of our fungal CRISPR-Cas9 system, we tested whether it could be used to mutagenize two different genes in *A. aculeatus*, 5-FOA counter-selectable *AACU_pyrG* and transcript ID 126899, which is homologous to *wA* in *A. nidulans* and to *albA* in *A. niger*. We therefore refer to 126899 as *AACU_albA* as mutation of this gene is expected to produce pale, rather than the wild-type black conidia. Accordingly, an *A. aculeatus* strain was transformed with CRISPR-Cas9 vectors encoding either no sgRNA, an sgRNA targeting Cas9 to the third exon of *AACU_albA* or an sgRNA targeting Cas9 to the second exon of *AACU_pyrG*. In all cases hygromycin was used for selection. As expected, transformants expressing only *cas9* formed black conidia (Fig 4A), which were unable to form colonies when plated on medium containing 5-FOA (Fig 4B). These results indicate that *AACU_albA* and *AACU_pyrG* were functional. In contrast, transformants expressing *cas9* and the specific *AACU_albA*-sgRNA gene readily formed white colonies, or colonies with white sectors (Fig 4C). Similarly, 5-FOA resistant colonies could easily be selected from a population of spores obtained from transformants expressing *cas9* and the specific *AACU_pyrG*-sgRNA gene (Fig 4D). We note that unlike mutagenesis of *yA* in *A. nidulans*, most colonies were solid white on the primary transformation plate suggesting that mutations occurred early after transformation. Like the *yA* mutations in *A. nidulans*, the *AACU_albA* mutations were genetically stable in the absence

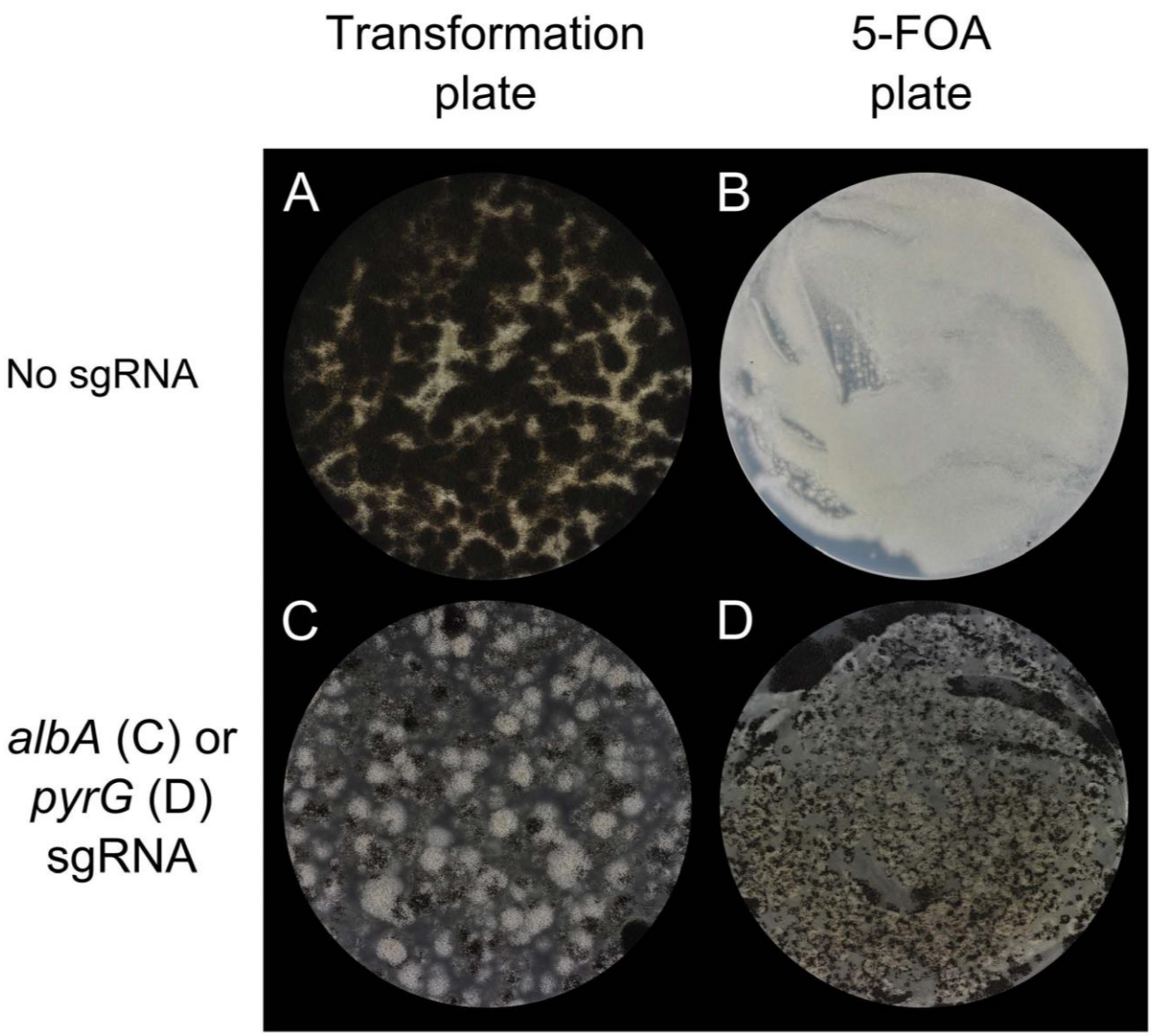

Fig 4. RNA guided Cas9 efficiently introduces directed mutations into *albA* and *pyrG* of *A. aculeatus*. *A. aculeatus* transformed with **A)** pFC332 encoding Cas9, but not a sgRNA plated on selective media. **B)** spores harvested from colonies shown in panel A, plated on medium containing 5-FOA, which impairs growth of *pyrG* wild-type strains preventing conidiation. **C)** pFC336 encoding Cas9 and an *albA* specific sgRNA. **D)** Spores harvested from transformants expressing Cas9 and pyrG specific sgRNA, pFC337, plated on medium containing 5-FOA. Note that colonies formed by *pyrG* mutant strains propagate normally and produce conidia.

doi:10.1371/journal.pone.0133085.g004

of the *AACU_albA*-CRISPR-Cas9 plasmid, as streak purified hygromycin sensitive strains produced only white conidia, see S1 Fig. Two white colonies and two 5-FOA resistant colonies were streak purified and the mutations were determined by sequencing PCR fragments covering the regions expected to be targeted by the *AACU_albA* and *AACU_pyrG* specific sgRNA. In all four cases small 1–10 bps deletions were identified at the sites expected to be cleaved by the Cas9 nuclease (Table 1). Hence, the typical RNA-guided Cas9 induced mutations in both

A. aculeatus and *A. nidulans* appear to be small deletions. Importantly, the fact that it is possible to specifically mutate genes in *A. nidulans* and *A. aculeatus*, despite that they belong to two different sections of *Aspergillus*, show that the basic components of our CRISPR-Cas9 system may work in quite distantly related filamentous fungi.

OPTIMuS, a script that facilitates construction of sgRNAs for targeting Cas9 to gene homologs in multiple species

The fact that the RNA guided Cas9 target site is composed of only 23 bps, i.e. the 3 bp PAM sequence and the adjacent 20 bps that match the protospacer (Fig 1A), prompted us to investigate whether it would be possible to design a protospacer, that is able to target several homologs of the same gene in different species. If so, mutation(s) could be engineered into homologs of the same gene in several species using the same vector construct. We therefore wrote a Perl script, OPTiMuS (One Protospacer for Targets in Multiple Species; see S1 OPTiMuS and S2 OPTiMUS) which facilitates identification of multi-species protospacers. Based on input FASTA files from selected species, the OPTiMuS identifies all PAM sequences in gene homologs and the accompanying protospacer sequences. Next OPTiMuS compares these protospacer-PAM sequences to identify those that are present in multiple species. Based on this analysis, the selected protospacer-PAM sequences are sorted after how many of the input sequences they may target. OPTiMuS can also be used to generate a list of protospacers, where all input sequences in a collection are matched at least once, hence, yielding a minimal set of protospacers needed to cover the collection. Additionally, as novel fully sequenced fungal genomes constantly emerge, we also included a feature in OPTiMuS that allows for a quick assessment of whether already constructed CRISPR-Cas9 vectors can be used for RNA-guided mutagenesis in the corresponding species. To test the functionality of OPTiMuS, we used AspGD [32] as a sequence resource to analyze *Aspergillus* homologs of *albA* for the presence of common protospacer-PAM sequences. The best hit was a protospacer-PAM sequence that potentially could target the gene homologs in eight species. Finally, since off-target Cas9 mutagenesis has been reported [16,17], we manually checked whether our selected protospacer-PAM sequence was prone for off-target effects. Specifically, we employed a simple BLAST search of the relevant genomes to look for sequences that match PAM and the 12 PAM proximal nucleotides of the protospacer [16,17]. Our selected *albA* protospacer-PAM sequence was not compromised by this test. Hence, it appears possible to generate sgRNAs that can target gene homologs in several different species without obvious off-target problems.

Introduction of specific mutations in *albA* homologs in five Aspergilli using a single protospacer sequence

A CRISPR-Cas9 vector encoding the *albA*-sgRNA identified by the OPTiMuS analysis was constructed. Since Cas9 was codon optimized for *A. niger*, and since we have in-house experience in transforming this species, we next decided to test the potential of the *albA*-sgRNA-Cas9 complex for its ability to induce mutagenesis in the five species, *A. brasiliensis* (ABRA), *A. carbonarius* (ACAR), *A. luchuensis* (ALUC), *A. niger* (ANIG) *and A. tubingensis* (ATUB), which all belong to section *Nigri*. Importantly, amongst these fungi, *A. brasiliensis* has to our knowledge not previously been genetically engineered. The strains were therefore transformed with the *albA* specific CRISPR-Cas9 vector, and its ability to stimulate specific RNA-guided Cas9 mutations was examined. Transformants were obtained for all species except *A. tubingensis*. With *A. brasiliensis* and *A. niger*, white transformants, or transformants containing white sectors, readily developed on the primary transformation plate (data not shown). For *A. carbonarius*, a few small white sectors developed amongst approximately 50 primary transformants and

for *A. luchuensis* a single white colony was obtained after restreaking four of the primary transformants. The single white *A. luchensis* transformant, two white colonies/sectors from *A. brasiliensis*, *A. carbonarius*, and *A. niger* were streak purified for further analysis. Like above, the region targeted by the *albA* specific Cas9 was PCR amplified and sequenced. In all five cases, indels, ranging from 7 bp to 83 bp were identified at the site expected to be cleaved by the Cas9 nuclease demonstrating that it is possible to generate CRISPR-Cas9 vectors with the ability to mutate homologs in several species. We note that despite that the two *A. niger albA* mutations were independently generated, our analysis showed that they were both identical 83 bp deletions. Perhaps a small three nucleotide ATA repeat present at the borders of the deletion promotes this type of event.

CRISPR-Cas9 induced DNA DSBs efficiently stimulate gene targeting in *A. nidulans* and in *A. aculeatu*

Next, we investigated whether DNA DSBs induced by the *yA* or *albA* specific sgRNA-Cas9 could stimulate gene targeting at these loci in *A. nidulans* and in *A. aculeatus*, respectively. We therefore co-transformed NHEJ proficient strains with linear (for *A. nidulans* and *A. aculeatus*) and circular gene-targeting substrates (for *A. aculeatus*) designed for deleting *yA* or *albA* in combination with either the *yA*-sgRNA- or *albA*-sgRNA-CRISPR-Cas9 vector (see Fig 1C and M&M). In these experiments, the gene-targeting substrates were composed by up- and downstream targeting sequences (approximately 2000 bp each) and by a selectable *Afl_pyrG* marker flanked by a direct repeat. Use of the *Afl_pyrG* for selection in *A. aculeatus* was possible due to the mutant *AACU-pyrG-1* strain created by RNA-guided mutagenesis above. The CRISPR-Cas9 plasmids were selected by an *AN_argB* marker in *A. nidulans* and by a hyg^R marker in *A. aculeatus*. Finally, linear and circular gene-targeting substrates were also co-transformed with a CRISPR-Cas9 negative control vector that does not encode the *yA* or *albA* specific sgRNAs.

The resulting transformants were green (*A. nidulans*) or black (*A. aculeatus*) if no RNA-guided Cas9 DNA DSBs were made at the *yA* or *albA* loci, respectively; and this was the case on plates selecting for the gene-targeting substrate (Fig 5A–5C) only, and on plates selecting for both the gene-targeting substrate and the empty CRISPR-Cas9 plasmid (Fig 5D–5F). The fact that no yellow or white mutant strains were observed likely reflects that gene targeting is inefficient in NHEJ proficient strains and that the numbers of transformants in the individual experiments are too low to produce rare transformants resulting from homologous recombination (HR).

In contrast, with the *A. nidulans* strain transformed with a *yA*-CRISPR-Cas9 vector and a linear *yA* gene-targeting substrate, approximately 90% of all transformants were solid yellow on both types of selection plates (Fig 5G and 5J). Pop-out recombinants, where the *AFL_pyrG* marker has been eliminated, were obtained on solid medium containing 5-FOA for eight yellow transformants from each plate. Further PCR analysis (see M&M) of these 16 strains by showed that they all contained a *yA* deletion. Four of the 16 transformants were examined in more detail to further validate that they contain a *yA* deletion. Specifically, we purified genomic DNA from two purified primary transformants and from two other transformants that were exposed to 5-FOA to eliminate the *AFL_pyrG* marker. Next, the four samples were subjected to Southern blotting using the direct repeat of the gene-targeting substrate as a probe. In agreement with the PCR analyses, the Southern blot revealed the bands that were expected if the *yA* locus has been correctly targeted in all four strains, see S2 Fig. Importantly, no other bands were observed strongly indicating that no additional copies of the gene-targeting substrate were present elsewhere in the genome.

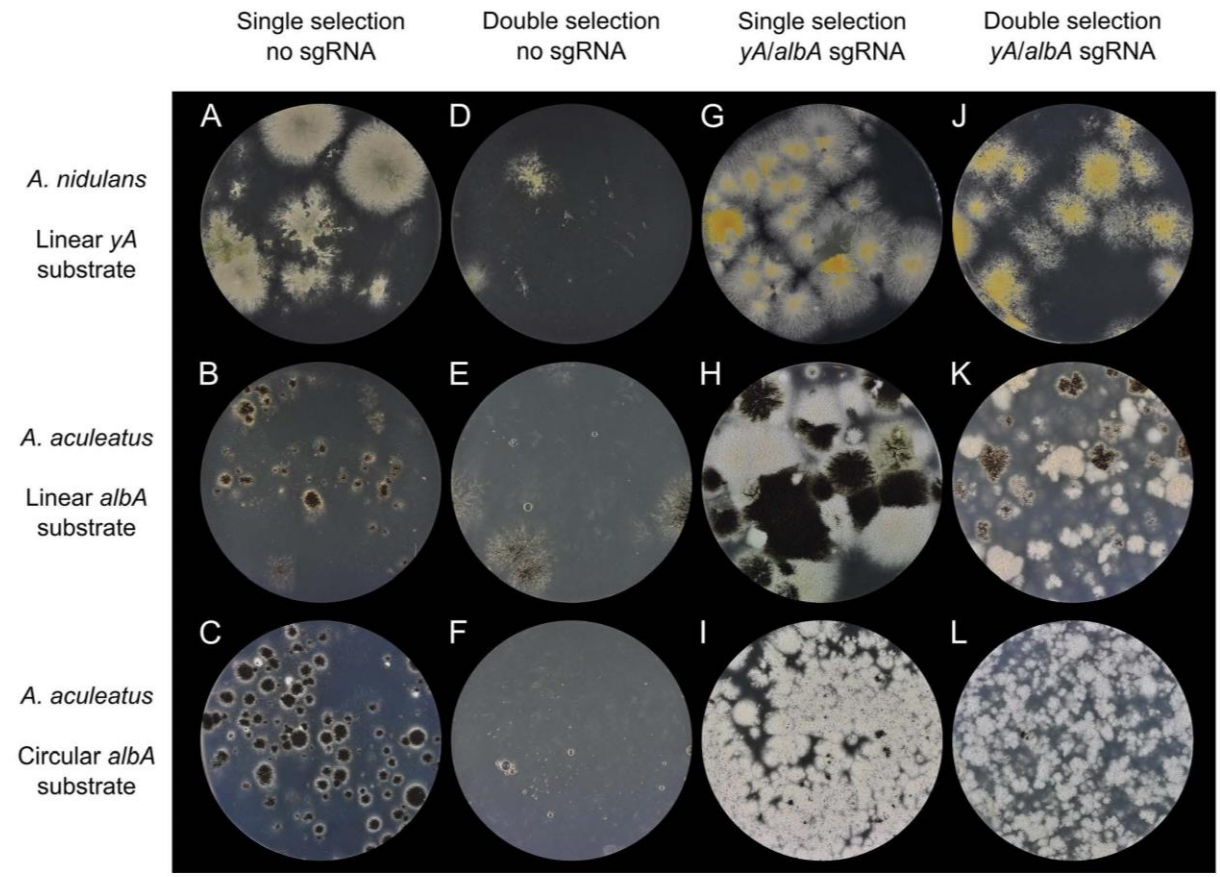

Fig 5. RNA guided induced DNA DSBs efficiently stimulate gene targeting in *A. nidulans* and *A. aculeatus*. *A. nidulans* and *A. aculeatus* were co-transformed with a CRISPR-Cas9 plasmid in combination with a gene targeting substrate. *yA* of *A. nidulans* was targeted by a linear gene targeting substrate; and *albA* of *A. aculeatus* was targeted by a linear as well as by a circular gene targeting substrate as indicated to the left of panels. The presence of an sgRNA gene (*yA* specific for *A. nidulans* and *albA* for *A. aculeatus*) gene in the CRISPR-Cas9-vector is indicated above panels. Selection conditions for each experiment are indicated above panels. Single selection refers to selection for the gene substrate alone; and double selection refers to selection for both the gene targeting substrate and the CRISPR-Cas9 vector.

doi:10.1371/journal.pone.0133085.g005

Next, we investigated whether CRISPR-Cas9 mediated gene targeting could be performed in *A. aculeatus*. An *A. aculeatus* strain was therefore transformed with an *albA*-CRISPR-Cas9 vector and linear or circular *albA* gene targeting substrate. The fact that most of the transformants were white independently of which type of plate was used for selection (Fig 5H,5I,5K and 5L) indicated that gene targeting was efficient. To confirm that white colonies represent *albA* deletions, two white transformants from each plate were streak purified and plated on medium containing 5-FOA. Pop-out recombinants were readily obtained in all four cases and PCR analysis of the resulting strains demonstrated that they all contained the *albA* deletion. Together these results demonstrate that gene targeting is greatly stimulated by CRISPR-Cas9 induced DNA DSB breaks at the *yA* locus in *A. nidulans* and at the *albA* locus in *A. aculeatus*.

Discussion

We have constructed a simple and versatile fungal CRISPR-Cas9 system and demonstrated its immediate potential by showing that (i) RNA-guided mutations can be introduced into specific alleles in six different species; (ii) that it is possible to use RNA-guided Cas9 induced DNA DSBs to stimulate otherwise inefficient gene targeting; and (iii) that RNA-guided mutagenesis and gene targeting can be used in succession. Importantly, the latter demonstrates that marker-free strains can quickly be engineered to allow for iterative gene targeting by e.g. mutating the endogenous *pyrG* and subsequently perform gene targeting using substrates based on *pyrG* (preferentially from a related species) as a recyclable marker. The only requirement for fungal RNA-guided Cas9 mutagenesis is the construction of target specific CRISPR-Cas9 vectors. In our version of the system this is readily achieved in a single USER-cloning based step, but other flexible cloning methods that allow for seamless joining of multiple fragments, e.g. by SLIC, Gibson DNA assembly, In-fusion or SLiCE [33–35], can easily substitute for USER-cloning if desirable. Importantly, the plasmid replicator (AMA1), the promoters, and terminators employed by the system have all been demonstrated to work in several fungal species. We therefore envision that the present version of our system, including using OPTiMuS to identify protospacers targeting the same gene, e.g. marker genes, in many different species, will speed up genetic engineering of newly sequenced fungi.

In one set of experiments, we showed that despite using an identical protospacer to stimulate RNA-guided mutagenesis, the efficiency of mutagenesis in the different species appeared to vary considerably. There may be several reasons why the mutation rate differs in individual species, but four deserve further discussion. Firstly, the *cas9* and sgRNA genes may be expressed differently in the different species due the use of fixed fungal promoters for their expression. If *cas9* and sgRNA genes are poorly expressed, RNA-guided mutagenesis may be inefficient. This problem could be alleviated by substituting the promoters in our current system by strong promoters from the host species. Similarly, the propagation stability of AMA1 based plasmids may also influence *cas9* and *sgRNA* gene expression levels, and this may also vary for the different species. In some cases, it may therefore be advantageous to integrate the *cas9*-sgRNA gene cassette into the genome of the host. Secondly, we have used a *cas9* gene that was codon optimized for efficient translation in *A. niger*. This may not be optimal for all species due to inefficient translation and folding of Cas9. We are currently constructing a GFP tagged version of Cas9 to quickly assess whether Cas9 levels in a new host is too low and needs to be optimized. Thirdly, since fungal protoplasts contain an unspecified number of nuclei, primary transformants are often heterokaryons. Subsequent isolation of pure homokaryotic mutant strains from the transformants is therefore frequently necessary. Pure strains are normally generated by restreaking conidia spores, but the efficiency of this process depends on the number of nuclei contained by the conidia; a number that varies from species to species. The fact that conidia from *A. nidulans* contain a single nucleus may explain why RNA-guided mutant strains are readily obtained in this species after one round of restreaking. In contrast, *A. carbonarius* and *A. luchuensis*, where the phenotypic development of RNA guided mutant strains required was less efficient, contain 2–5 nuclei [36]. Fourthly, the efficiency of accurate NHEJ repair may differ in individual species; and if the fidelity is high, the mutation rate will be low and depend on several cutting-sealing cycles, see Fig 1C, before the mutation is implemented. In this scenario, the mutations will develop and accumulate as a function of nuclear divisions.

We have investigated the use of RNA-guided Cas9 induced DNA DSBs to stimulate gene targeting. Traditionally gene targeting is performed using linear DNA substrates. Unlike intact circular DNA, linear DNA mimics DNA containing a DSB and will therefore recruit HR proteins to facilitate integration at a desired locus. However, linear DNA may also attract NHEJ

repair proteins, which will mediate undesired random integration of the DNA fragment into the genome. With most filamentous fungi, NHEJ is dominating HR for repair and for most wild-type fungi gene-targeting efficiencies are therefore low. The fact that CRISPR-Cas9 systems introduce DNA DSB at the target locus creates a different scenario. In this case, HR proteins recruited to the target locus will use the gene-targeting substrate as a template for repair, and it can therefore be delivered either as circular or as linear DNA. We find that circular gene-targeting substrates are much more efficient for gene targeting as compared to the corresponding linear substrates. This is likely explained by the fact that intact DNA circles do not integrate into the genome by the competing NHEJ pathway, hence, reducing the number of false positives.

Although, off-target effects after CRISPR-Cas9 mutagenesis may not be a major issue for filamentous fungi, it is advisable to keep *cas9* and sgRNA expression to a minimum. Importantly, for gene targeting, we show that it is possible to achieve high efficiency in experiments where selection is applied only to the gene-targeting substrate. Since AMA1 plasmids are readily lost without selection [37], we expect that the CRISPR-Cas9 plasmid was only present during cell divisions in the early phase of mycelium formation. Another way to minimize Cas9 activity would be to equip the *cas9* gene with an inducible promoter or to increase the specificity of Cas9 induced cleavage. The latter can be achieved by employing either shorter sgRNAs [38]; or mutant Cas9 that introduces nicks rather than DNA DSBs [39]. With the latter, DNA DSB formation for mutagenesis can be achieved if two mutant Cas9 species introduce nicks in close proximity on opposite strands of the DNA molecule. Higher specificity is achieved since binding of two different sgRNAs are necessary to target two Cas9 molecules to the locus. The two approaches have been successfully used to reduce off-target effects in several organisms [38–40].

In the present paper we have introduced mutations one by one into single loci in six different fungal species. However, for mammalian cells, and for yeast, it has been shown that gene editing of several loci can be performed simultaneously [8,9,20]. Moreover, nuclease inactive Cas9 variants have been fused (or not) to transcription factor activation domains; and in combination with an sgRNA, these Cas9 species may act as synthetic transcription factor activators or repressors; and we are currently investigating the possibility of implementing similar tools for filamentous fungi. For example, we envision that Cas9 activators can be very useful for analyzing genetically silent gene clusters for secondary metabolite production. For clusters containing a transcription factor (TF) gene, the cluster may be activated by an RNA-guided Cas9 transcription factor that turns this TF gene on. For clusters without a TF gene, activation of genes in the cluster could be activated by multiplex sgRNAs that targets Cas9 transcription factors to individual genes in the cluster. If the cluster genes are successfully and specifically activated in this manner, specific combinations of multiple sgRNAs and Cas9 could be used to

Table 2. Strains used in this study to implement functional CRISPR.

Species	Strain name	IBT[1] number	genotype
A. nidulans	NID5	27263	*argB2*, *pyrG89*, *veA1*
A. niger	ATCC 1015	28639	-
A. aculeatus	ATCC 16872	3244	-
A. luchuensis	CBS 106.47	32294	-
A. brasiliensis	CBS 101740	21946	-
A. carbonarius	ITEM 5010	31236	-

[1]IBT Culture Collection, www.bio.dtu.dk; contact Ulf Thrane.

doi:10.1371/journal.pone.0133085.t002

activate specific subsets of the genes in the cluster, which would greatly facilitate subsequent pathway elucidation. In conclusion, we show that CRISPR-Cas9 technology can be functionally adapted to work in filamentous fungi and we believe that the technology will be a major driver in forwarding the understanding of many aspects of fungal biology.

Materials and Methods

Strains and media

Escherichia coli strain DH5α was used to propagate all plasmids. The *Aspergillus* species used for implementation of CRISPR-Cas9 are listed in Table 2. Genomic DNA (gDNA) from fungal strains was isolated via FastDNA SPIN Kit for Soil DNA extraction kit (MP Biomedicals, USA). The mutant strains made in this study are listed in Table 1. All aspergilli were cultivated on standard solid glucose based minimal medium (MM) (1% glucose, 1x nitrate salt solution [41], 0.001% Thiamine, 1x trace metal solution [42], 2% agar), supplemented with 10 mM uridine (Uri), 10 mM uracil (Ura), and/or 4 mM L-arginine (Arg) when required. Solid plates containing 5-fluoroorotic acid (5-FOA) were made as MM+Arg+Uri+Ura supplemented with filter-sterilized 5-FOA (Sigma-Aldrich) to a final concentration of 1.3 mg/ml. For transformation media (TM) glucose was replaced with 1 M sucrose.

PCR, vector construction and protospacers

All PCR products for cloning purposes were amplified in 35 cycles using proof-reading PfuX7 polymerase [43], by touch-down PCR programs with maximum annealing temperature interval ranging from 68–59°C or 64–57°C. Standard reaction volumes were 50 μl including 1x Phusion HF Buffer (New England Biolabs, USA), 0.2 mm dNTPs, 0.4 μM primers (Integrated DNA Technologies (IDT), Belgium), 1 U PfuX7, <10 ng of gDNA, 3% DMSO. PfuX7 can be substituted by Phusion U (Life Technologies). All vectors were assembled by USER cloning or USER fusion as described previously [44,45]. Vectors were constructed by USER cloning using plasmid backbones previously presented by [45]; for details concerning vector construction, see Fig 2, S1 Protocol, and S1 Table. The *cas9* gene encoding *Streptococcus pyogenes* Cas9 was codon optimized for translation in *A. niger* and synthetized by GeneScript in two parts, for its sequence, see genbank accession number KT031982. Sequences for the four vectors pFC330-333, and the gBlock encoding the *yA* sgRNA expression cassette, defining our present fungal CRISPR-Cas9 platform can be found at genbank with the accession numbers KT031983, KT031984, KT031985, KT031986 and KT031987. All vectors are available on request. Protospacers used to target individual fungal genes are presented in Table 3.

Table 3. Protospacers used in this study.

Protospacer sequence	Gene	Species
GGCGGAGTATCATAACATCG	*yA*	*A. nidulans*
AGTGGGATCTCAAGAACTAC	*albA*	*A. niger*
AGTGGGATCTCAAGAACTAC	*albA*	*A. brasiliensis*
AGTGGGATCTCAAGAACTAC	*albA*	*A. luchuensis*
AGTGGGATCTCAAGAACTAC	*albA*	*A. carbonarius*
GGCGACATGTTGAAGAACCG	*albA*	*A. aculeatus*
GATGCTGCAGTTGATGATGT	*pyrG*	*A. aculeatus*

doi:10.1371/journal.pone.0133085.t003

Transformation and strain validation by Tissue-PCR

Protoplastation were performed as described by Nielsen et al 2006 [31]. For transformation using either *pyrG* or *argB* as genetic marker, 10^7 protoplasts and ~3 μg of digested DNA and 150 μl PCT solution were incubated for 10 min at room temperature, followed by adding of 250 μl ATB plating on 1 M sucrose based TM with selection. All TM plates were incubated at 30°C, except for *A. nidulans* transformation at 37°C. For transformation utilizing hygromycin selection, 10^7 protoplasts and ~3 μg of digested DNA were incubated at 15 min on ice, then 1 mL PCT was added and the mix incubated for 15 min at room temperature. 100 μg/ml hygromycin B (Invivogene, USA) was added to 15 ml molten 1 M sorbitol based TM (TMsh; ~45°C), and immediately poured into an empty 9 cm petri dish. After 24 h incubation at 30°C, an overlay of 15 ml TMsh was added. All candidate transformants were streak purified prior to verification.

All strains were verified by tissue-PCR analysis using mycelium as the source of DNA. For the specific PCR protocol, see S2 Protocol, and for primers refer to S1 Table. Primers for gene-deletion analysis were designed to bind up- and downstream outside the region eliminated by the gene-targeting substrate. This setup detects wild-type sequences present in transformants that were for false positive and heterokaryons as well as for gene-deletion strains derived after marker elimination by direct repeat recombination. Transformants where a deleted gene was replaced by a marker gene was validated in two PCR reactions using a primer pair where one primer binds outside the upstream targeting region and the other binds inside the marker gene; and a pair where one of the primes bind inside the marker gene and the other binds outside of the downstream targeting region. For the experimental design and details of protocol, see S2 Protocol. Wild-type gDNA, as well as wild-type mycelium, was always included as controls to benchmark tissue-PCR efficiency for wild-type reaction success. All primers are listed in S1 Table.

Note added in proof

While this manuscript was in revision, a functional CRISPR-Cas9 system for gene editing of *Trichoderma reesei* was published [46] demonstrating that this method may work more widely in filamentous fungi.

Supporting Information

S1 Fig. White *A. aculeatus* mutant strains propagate in a phenotypically stable manner.
(PPTX)

S2 Fig. Southern blot analysis of the *yA* locus after CRISPR-Cas9 mediated gene targeting.
(PPTX)

S1 OPTiMuS. Users guide to OPTiMuS.
(DOCX)

S2 OPTiMuS. Perl script file for OPTiMuS.
(PL)

S1 Protocol. Vector construction.
(DOCX)

S2 Protocol. PCR analysis and protocol.
(DOCX)

S1 Table. Primer used in this study.
(DOCX)

Acknowledgments

We thank Morten Kielland-Brandt for critically reading this manuscript; and Peter Wad Sackett, Benjamin Bennike Aagren and Tobias Bennike Aagren for technical help and inspiration during the development of OPTIMuS; and finally, Magdalena Skrzypczak for help with protoplastation of the different fungal species. This work was supported by grant 11–116803 from The Danish Council for Strategic Research.

Author Contributions

Conceived and designed the experiments: CSN JBN UHM. Performed the experiments: CSN JBN MEK. Analyzed the data: CSN JBN MEK. Contributed reagents/materials/analysis tools: CSN JBN MEK. Wrote the paper: CSN JBN UHM.

References

1. Fleissner A, Dersch P. Expression and export: Recombinant protein production systems for Aspergillus. Appl Microbiol Biotechnol. 2010; 87: 1255–1270. doi: 10.1007/s00253-010-2672-6 PMID: 20532762
2. Moretti a., Susca a., Mulé G, Logrieco a. F, Proctor RH. Molecular biodiversity of mycotoxigenic fungi that threaten food safety. Int J Food Microbiol. Elsevier B.V.; 2013; 167: 57–66. doi: 10.1016/j.ijfoodmicro.2013.06.033 PMID: 23859402
3. Wu F, Groopman JD, Pestka JJ. Public health impacts of foodborne mycotoxins. Annu Rev Food Sci Technol. 2014; 5: 351–372. doi: 10.1146/annurev-food-030713-092431 PMID: 24422587
4. Ballance DJ, Buxton FP, Turner G. Transformation of Aspergillus nidulans by the orotidine-5´-phosphate decarboxylase gene of Neurospora crassa. Biochmical Biophys Res Commun. 1983; 112: 284–289.
5. Boeke JD, La Croute F, Fink GR. A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. MGG Mol Gen Genet. 1984; 197: 345–346. doi: 10.1007/BF00330984
6. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. Elife. 2013; 2013: e00471. doi: 10.7554/eLife.00471
7. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol. Nature Publishing Group; 2013; 31: 227–9. doi: 10.1038/nbt.2501 PMID: 23360964
8. Cong L, Ran F, Cox D, Lin S, Barretto R. Multiplex Genome Engineering Using CRISPR / Cas Systems. Science (80-). 2013; 339: 819–23. doi: 10.1038/nbt1319 PMID: 23287718
9. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-Guided Human Genome Engineering via Cas9_Sup. Science (80-). 2013; 339: 823–6. doi: 10.1126/science.1232033 PMID: 23287722
10. Cho SW, Kim S, Kim JM, Kim J-S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol. Nature Publishing Group; 2013; 31: 230–2. doi: 10.1038/nbt.2507 PMID: 23360966
11. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science (80-). 2012; 337: 816–21. doi: 10.1126/science.1225829 PMID: 22745249
12. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A. 1996; 93: 1156–1160. doi: 10.1073/pnas.93.3.1156 PMID: 8577732
13. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. Science. 2009; 326: 1509–1512. doi: 10.1126/science.1178811 PMID: 19933107
14. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science. 2009; 326: 1501. doi: 10.1126/science.1178817 PMID: 19933106
15. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol. Nature Publishing Group; 2014; 32: 347–55. doi: 10.1038/nbt.2842 PMID: 24584096

16. Hsu PD, Scott D a, Weinstein J a, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol. 2013; 31: 827–32. doi: 10.1038/nbt.2647 PMID: 23873081

17. Cradick TJ, Fine EJ, Antico CJ, Bao G. CRISPR/Cas9 systems targeting ??-globin and CCR5 genes have substantial off-target activity. Nucleic Acids Res. 2013; 41: 9584–9592. doi: 10.1093/nar/gkt714 PMID: 23939622

18. Jakoči nas T, Bonde I, Herrgård M, Harrison SJ, Kristensen M, Pedersen LE, et al. Multiplex metabolic pathway engineering using CRISPR/Cas9 in Saccharomyces cerevisiae. Metab Eng. 2015; 28: 213–222. doi: 10.1016/j.ymben.2015.01.008 PMID: 25638686

19. Ryan OW, Skerker JM, Maurer MJ, Li X, Tsai JC, Poddar S, et al. Selection of chromosomal DNA libraries using a multiplex CRISPR system. Elife. 2014; e03703. doi: 10.7554/eLife.03703

20. Dicarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. 2013; 41: 4336–4343. doi: 10.1093/nar/gkt135 PMID: 23460208

21. Belhaj K, Chaparro-Garcia A, Kamoun S, Nekrasov V. Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. Plant Methods. Plant Methods; 2013; 9: 39. doi: 10.1186/1746-4811-9-39 PMID: 24112467

22. Gems D, Johnstone IL, Clutterbuck AJ. An autonomously replicating plasmid transforms Aspergillus nidulans at high frequency. Gene. 1991; 98: 61–67. doi: 10.1016/0378-1119(91)90104-J PMID: 2013411

23. Ruiz-Díez B. A Review: Strategies for the transformation of filamentous fungi. Journal of Applied Microbiology. 2002. pp. 189–195. doi: 10.1046/j.1365-2672.2002.01516.x PMID: 11849345

24. Kitamoto N, Matsui J, Kawai Y, Kato A, Yoshino S, Ohmiya K, et al. Utilization of the TEF1-α gene (TEF1) promoter for expression of polygalacturonase genes, pgaA and pgaB, in Aspergillus oryzae. Appl Microbiol Biotechnol. 1998; 50: 85–92. doi: 10.1007/s002530051260 PMID: 9720204

25. Tamano K, Bruno KS, Karagiosis SA, Culley DE, Deng S, Collett JR, et al. Increased production of fatty acids and triglycerides in Aspergillus oryzae by enhancing expressions of fatty acid synthesis-related genes. Appl Environ Microbiol. 2013; 269–281. doi: 10.1007/s00253-012-4193-y

26. Kozak M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell. 1986; 44: 283–292. doi: 10.1016/0092-8674(86)90762-2 PMID: 3943125

27. Gao Y, Zhao Y. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. J Integr Plant Biol. 2014; 56: 343–349. doi: 10.1111/jipb.12152 PMID: 24373158

28. Jacobs JZ, Ciccaglione KM, Tournier V, Zaratiegui M. Implementation of the CRISPR-Cas9 system in fission yeast. Nat Commun. Nature Publishing Group; 2014; 5: 5344. doi: 10.1038/ncomms6344 PMID: 25352017

29. Nissim L, Perli SD, Fridkin A, Perez-pinera P, Lu TK. Multiplexed and Programmable Regulation of Gene Networks with an Integrated RNA and CRISPR / Cas Toolkit in Human Cells. 2014;

30. O'Hara EB, Timberlake WE. Molecular characterization of the Aspergillus nidulans yA locus. Genetics. 1989; 121: 249–254. PMID: 2659435

31. Nielsen ML, Albertsen L, Lettier G, Nielsen JB, Mortensen UVH. Efficient PCR-based gene targeting with a recyclable marker for Aspergillus nidulans. Fungal Genet Biol. 2006; 43: 54–64. doi: 10.1016/j.fgb.2005.09.005 PMID: 16289954

32. Arnaud MB, Cerqueira GC, Inglis DO, Skrzypek MS, Binkley J, Chibucos MC, et al. The Aspergillus Genome Database (AspGD): Recent developments in comprehensive multispecies curation, comparative genomics and community resources. Nucleic Acids Res. 2012; 40: D653–9. doi: 10.1093/nar/gkr875 PMID: 22080559

33. Gibson DG, Young L, Chuang R, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods. 2009; 6: 343–345. doi: 10.1038/nmeth.1318 PMID: 19363495

34. Li MZ, Elledge SJ. SLIC: A method for sequence- and ligation-independent cloning. Methods Mol Biol. 2012; 852: 51–59. doi: 10.1007/978-1-61779-564-0_5 PMID: 22328425

35. Zhang Y, Werling U, Edelmann W. SLiCE: A novel bacterial cell extract-based DNA cloning method. Nucleic Acids Res. 2012; 40: 1–10. doi: 10.1093/nar/gkr1288 PMID: 21908400

36. Yuill E. The numbers of nuclei in conidia of aspergilli. Transactions of the British Mycological Society. 1950. pp. 324–IN10. doi: 10.1016/S0007-1536(50)80083-9

37. Aleksenko A, Clutterbuck AJ. Autonomous plasmid replication in Aspergillus nidulans: AMA1 and MATE elements. Fungal Genet Biol. 1997; 21: 373–387. doi: 10.1006/fgbi.1997.0980 PMID: 9290250

38. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat Biotechnol. 2014; 32: 279–84. doi: 10.1038/nbt.2808 PMID: 24463574

39. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, et al. Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity. Cell. Elsevier; 2013; 154: 1380–1389. doi: 10.1016/j.cell.2013.08.021 PMID: 23992846

40. Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol. 2013; 31: 833–8. doi: 10.1038/nbt.2675 PMID: 23907171

41. Kaminskyj SGW. Fundamentals of growoth, storage, genetics and microscopy of Aspergillus nidulans. Fungal Genet Newsl. 2001; 25–31.

42. Cove DJ. The induction and repression of nitrate reductase in the fungus Aspergillus nidulans. Biochim Biophys Acta. Elsevier Publishing Company, Amsterdam; 1966; 113: 51–56. doi: 10.1016/S0926-6593(66)80120-0 PMID: 5940632

43. Nørholm MHH. A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. BMC Biotechnol. 2010; 10: 21. doi: 10.1186/1472-6750-10-21 PMID: 20233396

44. Nour-Eldin HH, Geu-Flores F, Halkier BA. USER cloning and USER fusion: the ideal cloning techniques for small and big laboratories. Methods Mol Biol. 2010; 643: 185–200. doi: 10.1007/978-1-60761-723-5 PMID: 20552452

45. Hansen BG, Salomonsen B, Nielsen MT, Nielsen JB, Hansen NB, Nielsen KF, et al. Versatile enzyme expression and characterization system for Aspergillus nidulans, with the Penicillium brevicompactum polyketide synthase gene from the mycophenolic acid gene cluster as a test case. Appl Environ Microbiol. 2011; 77: 3044–3051. doi: 10.1128/AEM.01768-10 PMID: 21398493

46. Liu R, Chen L, Jiang Y, Zhou Z, Zou G. Efficient genome editing in filamentous fungus Trichoderma reesei using the CRISPR/Cas9 system. Cell Discov. 2015; 1: 15007. doi: 10.1038/celldisc.2015.7

Chapter 4: Expanding tools and applications of the CRISPR-Cas9 system in *Aspergillus nidulans*

Introduction

The genome editing technology CRISPR-Cas9 has recently been adapted for the use in various filamentous fungi [1–4]. So far, the focus has mainly been on finding a setup that allows for the bacterial/archaeal RNA-guided nuclease to be successfully used in these organisms and how to use it to either introduce mutations by faulty non-homologous end-joining (NHEJ) or by boosting frequencies repair by homologous recombination (HR). However in other cell systems such as *S. cerevisiae* or mammalian cell systems the CRISPR-Cas9 technology has been a useful tool since 2013, and much more development has been done.

One of the main concerns when using genome editing technologies such as the CRISPR-Cas9 technology is the risk of creating unintended double-strand-breaks (DSB). Of course, this is a problem that scales with genome size, since specificity is determined by 20 bp. The bigger the genome is, the harder is it to find 20 unique bases in the desired location. There have been lots of efforts to come up with strategies to limit off-targeting and to gauge the scope of the problem. Strategies to limit off-targeting includes inactivating one of the sites in Cas9 responsible for cleaving one of the DNA strands, creating a nickase [5,6], essentially extending the target site when using two sgRNAs causing nicks in close proximity on opposite DNA strand. Similar but more elaborate is a strategy where both domains responsible for cleavage is inactivated, and the Cas9 protein is then fused to a dimeric FokI nuclease [7,8]. The main differences between the paired nickases and Cas9-FokI fusions are that a single nickase can cause a nick by itself, and the distance between two nicks can be somewhat flexible, being able to induce HR or NHEJ with distances up to approximately 100 bp [5,6,9]. In comparison the Cas9-FokI fusion only induces a DSB when two monomers binds in close proximity, 10-25 bp apart [7,8]. A third strategy is based on truncating the protospacer down to 16-19 bp instead of the normal 20 bp, the argument being that with fewer based to provide basepairing, the lower the tolerance is for mismatches [10]. All three strategies have been shown to be able to reduce off-targeting effects.

Interestingly, to which degree off-targeting is a problem is still up for discussion. While some studies report off-targeting to be a problem [11,12], others do not. Two recent studies included deep sequencing of either *S. cerevisiae* or CHO cells at predicted off-targeting sites [13,14] showed no indel formation attributable to off-targeting of Cas9. Since the verdict is still out whether off-targeting is a problem in general and specifically in Aspergilli, we here test two methods for reducing off-targeting.

One of the exciting feats done with CRISPR-Cas9 in both *S. cerevisiae* and CHO cells is the ability to perform various genetic modifications without the need for integration of a selection marker and several modifications at a time [14–16]. In Aspergilli this ability would be extremely valuable, since available selection markers are often limited and markers such as *pyrG*, which can be recycled by direct repeat recombination [17,18], are even more limited. Each cycle of protoplastation, transformation, verification and marker recycling is time consuming, and if multiple modifications can be done at once, it will greatly accelerate the rate of which genetic engineering can be performed. To my knowledge, instances of doing multiple genetic modifications at once, outside ingration of multiple copies inserted by NHEJ, without the use of genome editing technologies, have not been published. An especially interesting use for genetic modifications without an accompanying selection marker is for the introduction of point mutations and other very small modifications, such as introduction of small tags etc. The ability to introduce specific point mutations are very powerful, for purposes such as protein engineering, metabolic engineering, studies in protein function etc., however it is also one of the most difficult types of genetic changes to introduce, since unlike gene deletions or insertions, it should preferably to be scar-less except for the desired change. While in some cases, co-insertion of an adjacent marker gene has to be accepted, this risks introducing positional effects on transcription levels of the gene of interest. In yeast several methods have been developed for seamless and markerless introduction of point mutations, such as perfetto delitto [19] or the 50:50 [20]. Both methods takes advantage of the fact the *S. cerevisiae* can integrate DNA with high frequencies of HR, even with very short flanks of homology, like 40-50 bp [21] which is short enough to include in a primer tail. Both methods are based on using a counter-selectable marker. While the counter-selectable markers are available in many fungal species, the requirement for flank length for homologous recombination is much higher, typically around 500-1000 even if non-homologous end-joining is impaired [22]. In fungal systems where counter-selectable markers are available point mutations can be introduced seamlessly [23], though the process is time-consuming, requiring complex vector constructions. In

filamentous fungi where counter-selectable markers are not available, it is a lot more complicated and usually leaves a selection marker [24]. However, with the recent adaptation of CRISPR-Cas9 for use in filamentous fungi, it should be possible to enhance the processes for doing many types of genetic engineering. In this study we demonstrate how CRISPR-Cas9 can be used in *Aspergillus nidulans* to do various genetic modification without the introduction of a selection marker and also how single-stranded DNA (ssDNA) oligos can be used as a repair template for HR, as an alternative to conventional gene targeting substrates.

Results

Testing different promoters for the ability to express gRNAs

Besides the ribozyme based method for expression of the gRNA presented in chapter three, several other methods were also tested. One method tested was just to express the gRNA from a conventional RNA polymerase II promoter, such as those used for protein expression. The gRNA with a protospacer targeting the *yA* [3] was inserted directly into the *A. nidulans gpdA* promoter and the *trpC* terminator and inserted into the pFC331 vector, containing Cas9 expressed from the *A. nidulans tef1* promoter, an *argB* marker gene and the AMA1 element.

In both plants and mammalian systems, a common way to express the gRNA is from the U6 promoter, which is transcribed by the RNA polymerase III, and thus avoids polyadenylation and export from the nucleus. The AspGD database [25] was searched for U6 small nuclear RNA, and three hits came up, annotated as U6-1, U6-2 and U6-3 in *Aspergillus fumigatus*. For promoter and terminator respectively, 307 bp upstream and 53 downstream of the U6-1 were used to express the gRNA with the *yA* protospacer. This construct was cloned into pFC331.

Similar to the U6 promoter, the U3 promoter has been successfully used in some plants to express sgRNAs [26]. Using AspGD a single U3 snoRNA was found, from *A. fumigatus*. 486 bp upstream of the U3 snoRNA was used as promoter combined with the *yA* specific sgRNA and the SUP4 terminator and the construct was inserted into pFC331.

Another alternative used in *Saccharomyces cerevisiae* is the *SNR52* promoter, since the yeast U6 promoter, unlike that of the U6 in vertebrae is dependent on downstream elements as well as upstream elements [27] which limits its use for heterologous expression. To test if the *SNR52*

promoter could be used to express gRNAs in Aspergilli, a construct consisting of the *S. cerevisiae SNR52* promoter, the sgRNA with the *yA* protospacer and the *S. cerevisiae SUP4* terminator was cloned into pFC331

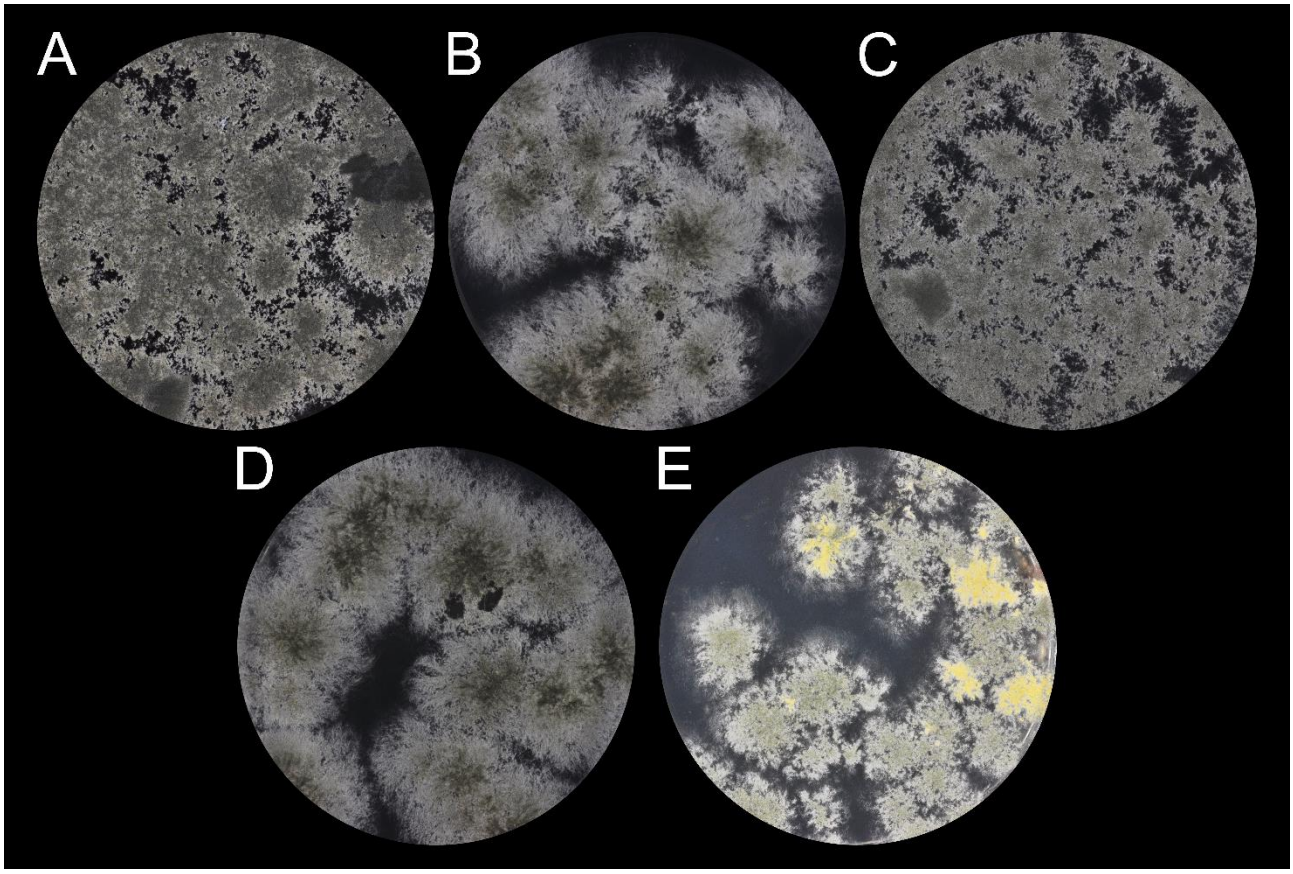

Figure 4.1 Test of suitable promoter for sgRNA expression. All the transformants are based on NID5, all the vectors are based on pFC331 and the sgRNA is the same, the only difference is the promoter and terminator used for sgRNA expression. A) *A. nidulans gpdA* promoter and *trpC* terminator. B) *S. cerevisiae SNR52* promoter and *SUP4* terminator. C) *A. fumigatus U6-1* promoter and terminator. D) *A. fumigatus U3* promoter and *S. cerevisiae SUP4* terminator. E) *A. nidulans gpdA* promoter and *trpC* terminator, but with the yA specific sgRNA flanked by a hammerhead and a HDV ribozyme.

The different vectors were all transformed into NID5, which is NHEJ profient, but is auxotrophic for argine, uridine and uracil, and the results are summarized into Table 1. For comparison a transformation with pFC334 [3], which contains the same *yA* protospacer, and the *gpdA* promoter and *trpC* terminator, but with the sgRNA flanked by ribozymes, was included. The plates can be seen in Figure 4.1, panel A-E show sgRNA expression from the *A. nidula*ns *gpdA* promoter the

yeast *SNR52* promoter, the *A. fumigatus* U6-1 promoter, the *A. fumigatus* U3 promoter and the *A. nidulans gpdA* promoter but flanked with two ribozymes[3], respectively.

Table 4.1 Test of different promoters for sgRNA expression, measured in their ability to induce mutagenesis in the *yA* gene leading to yellow-spored colonies

Promoter	Yellow or partial yellow /Green (frequency)
P_{gpdA}	0/95 (0%)
P_{SNR52}	0/55 (0%)
$P_{U6\text{-}1}$	0/64 (0%)
P_{U3}	0/57 (0%)
P_{gpdA} (ribozymes)	15/54 (22%)

Testing strategies for limiting off-targeting

One of the main concerns for using the CRISPR-Cas9 is the risk of off-targeting effects, in the form of unintended DSB (double strand breaks) in other places of the genome besides the targeted site. One strategy for limiting off-targeting is shortening the protospacer, since with a shorter protospacer, the requirements for specificity are higher for getting sufficient binding between gRNA and chromosomal DNA [10]. A construct was designed containing the same protospacer as pFC334, except shortened with 3 bp in the PAM distal end, for a total of 17 bp and cloned into pFC330 containing the *pyrG* selection marker, Cas9 and the AMA1 element. The resulting vector pCRISPR1-14p was transformed into NID5 and the resulting colonies displayed two distinct phenotypes, as seen in Figure 4.2, panel A and B. One was small and sporeless, the other type had no spores either, but an aberrant morphology with the texture of the colony raised. No spore color was discernable in either type. Sequencing of the region showed that the sequence was correct. Therefore, while shortening protospacers may be a viable strategy in human cells, it does not seem to be so in *Aspergillus nidulans*.

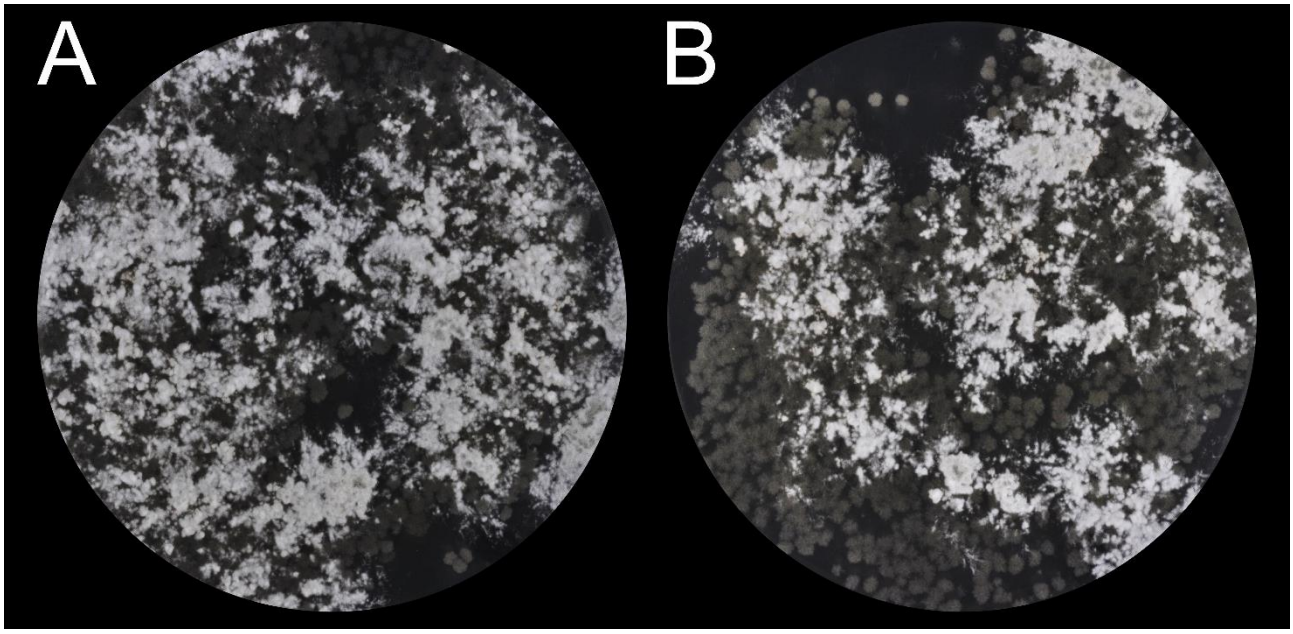

Figure 4.2 Test of shortened protospacer. A) and B) both shows transformation plates of NID5 transformed with a vector with a sgRNA in which the protospacer is only 17 bp rather than 20, but otherwise the same as the one in pFC334.

Another method used to limit off-targeting is to disable one of the domains in the Cas9 protein, since two domains are both responsible for cutting one strand of the targeted DNA, thus by disabling one of them, the protein gets turned into a targeted nickase instead. Combined with two sgRNAs to make two nicks in close proximity on opposite DNA strands, this strategy has been successfully used to limit off-targeting in both plants and mammalian cell lines [5,28]. Four vectors were constructed, pCas9-D10A-argB, pCas9-H840A-argB, pCas9-D10A-pyrG and pCas9-H840A-pyrG, similar to pFC330 and pFC331 except that the Cas9 carries either a D10A or H840A mutation, disrupting one of the cleaving domains, turning it into a nickase. Next two sets of protospacers were designed to make two nicks at opposite strands with a distance of either 16 bp or 65 bp. The protospacers were expressed from the *H3-H4* bidirectional histone promoter, instead of the *gpdA* promoter which has been used so far and cloned into the argB containing markers. The four resulting plasmids, two different mutations and two different protospacer configurations, all with the *argB* selection marker, were transformed into NID5. Due to low pH in the transformation media, some of the spores looked quite yellow, but a different kind of yellow than expected from a *yA* disruption, as seen in Figure 4.3 panel A-D. Four of the most intensely yellow looking colonies from each plate were re-streaked to MM-uri/ura, however all of them turned out to be green. Next the two vectors with the D10A nickase and either protospacer configuration inserted, was co-transformed with the *yA* deletion vector [3], p6f-yA, linearized or circular in NID5, selecting for

pyrG. Figure 3, panel E and F show transformation plates with the vector having 16 bp overhang with linear and circular p6f-yA, respectively. Figure 4.3, panel G and H show transformation plates with the vector having 65 bp overhang with linear and circular p6f-yA, respectively. All cases results in the two types of colonies, a small and sporeless one, and normal looking ones. In general there was very few normal ones, but of among those, yellow spores was the most common type, indicating the *yA* deletion had been successful. Furthermore the most normal colonies resulted from transformation in which a circular gene targeting substrate and short overhangs of 16 bp were used, suggesting that using a circular substrate and relatively short overhangs are prefered.

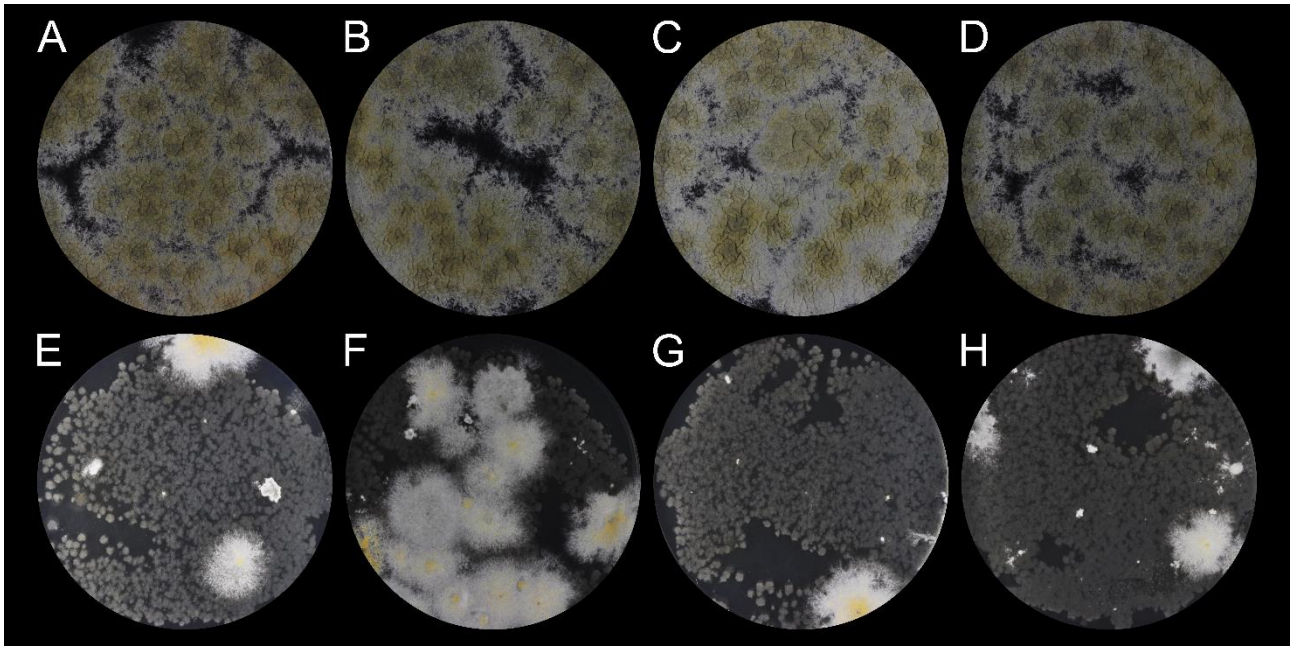

Figure 4.3 Test of Cas9 nickases. The nickase vectors contains either a D10A or an H840A mutation and two protospacer creating overhangs of either 16bp or 65 bp. A)-D) are transformed with a nickase vector alone, while E) and G) are co-transformed with a linear vector p6f-yA for deletion of yA, while F) and H) are co-transformed with circular p6f-yA. A) D10A, 16 bp. B) D10A, 65 bp. C) H840A, 16 bp. D) H840A, 65 bp. E) and F) D10A, 16 bp. G) and H) D10A and 65 bp.

Investigating the effect of inducing a double strand break in an NHEJ deficient strain.

So far, one of the main applications of the CRISPR-Cas9 technology in filamentous fungi has been to use it to boost gene targeting frequencies in a NHEJ proficient background. However, it is common to use NHEJ deficient strains for genetic engineering purposes and it would be useful to know how such a strain interacts with the DSB inducing properties of CRISPR-Cas9. The strain

NID1, identical to NID5 except it also contains an *nkuA* deletion, was transformed with pFC334. Two transformation plates can be seen in Figure 4.4, panel A and B. Out of approximately 88 colonies not a single one displayed a yellow phenotype. While it is difficult to compare total number of colonies with the similar transformations in NID5, due to batch variation in protoplasts, the general tendency was that NID1 yielded fewer colonies than NID5. Considering that there are no template available, except for certain stages of the cell cycle, repair by HR is unlikely. With *nkuA* deleted, NHEJ is not available and repair by MMEJ would result in errors and as such result in the yellow phenotype. A likely scenario is that those colonies which in a NHEJ proficient strain would turn yellow due to faulty repair by NHEJ, dies in an NHEJ deficient background. However due to the multicellular nature some nuclei escapes Cas9, resulting in the green colonies observed.

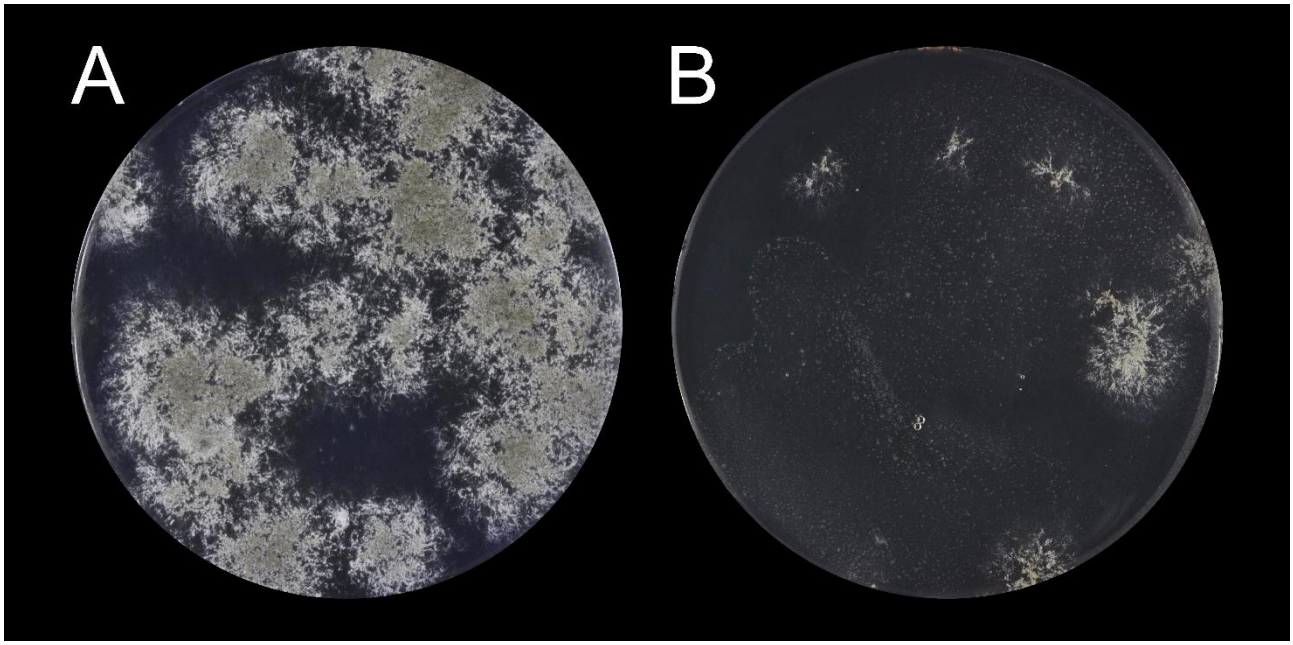

Figure 4.4 Using CRISPR-Cas9 in an NHEJ deficient background. A) and B) both shows transformation plates of NID1 transformed with pFC334

Expanding CRISPR-Cas9 applications to avoid integration of selection marker

In Chapter 3 CRISPR-Cas9 was used to assist integration by homologous recombination in which a gene targeting substrate including a selection marker was integrated into the genome. In this case, selection for the vector carrying Cas9 and the sgRNA was not necessary. However if this scenario could be flipped so selection for the repair template could be avoided, it would be extremely beneficial, since it would allow for seamless integrations, deletions and point mutations, avoiding

potential positional effects from a neighboring marker gene and it would also allow for multiplexing, which is otherwise limited by the availability of selection markers. To test this I transformed NID5 with pFC334 and, either a linearized or a circular *yA* deletion vector, p6f-yA, but selecting for *argB* instead of *pyrG*. Most of the colony mass was green, but many contained flecks of yellow. Unlike when selecting for a marker gene to be integrated, the rate of yellow spores, suggesting an integration event, was significantly lower , as seen on Figure 4.5, panel A and B, which used a linear and a circular repair template respectively.

When not selecting for the DNA to be integrated, there is the possibility that a DSB happens but no template is available and therefore the break is repaired NHEJ, faulty or not. As such, it seems reasonable that disabling NHEJ, as it is done for improving classical gene targeting, would also enhance integration frequencies in this scenario. Especially considering that NID1 transformed with pFC334 did not yield a single yellow colony. To test if this was the case, the experiment was repeated in NID1, and transformation plates can be seen in Figure 4.5, panel C and D. To assess whether the yellow phenotype indeed was a result of the deletion template being integrated and not just mutagenesis by faulty NHEJ, 10 colonies from each setup, NID1 or NID5, linear or circular substrate, was stabbed to a MM plate containing arginine but neither uridine or uracil. For NID5 transformed with a linear gene targeting substrate, 7 out of 10 grew. For NID5 transformed with circular gene targeting substrate 9 out of 10, while for NID1 transformed with either kind of substrate 10 out of 10 grew, confirming integration of the substrate. From the plates it can be seen that the tendency is that using a circular substrate seem more efficient that a linear substrate. Even greater impact however, is the effect of disabling NHEJ.

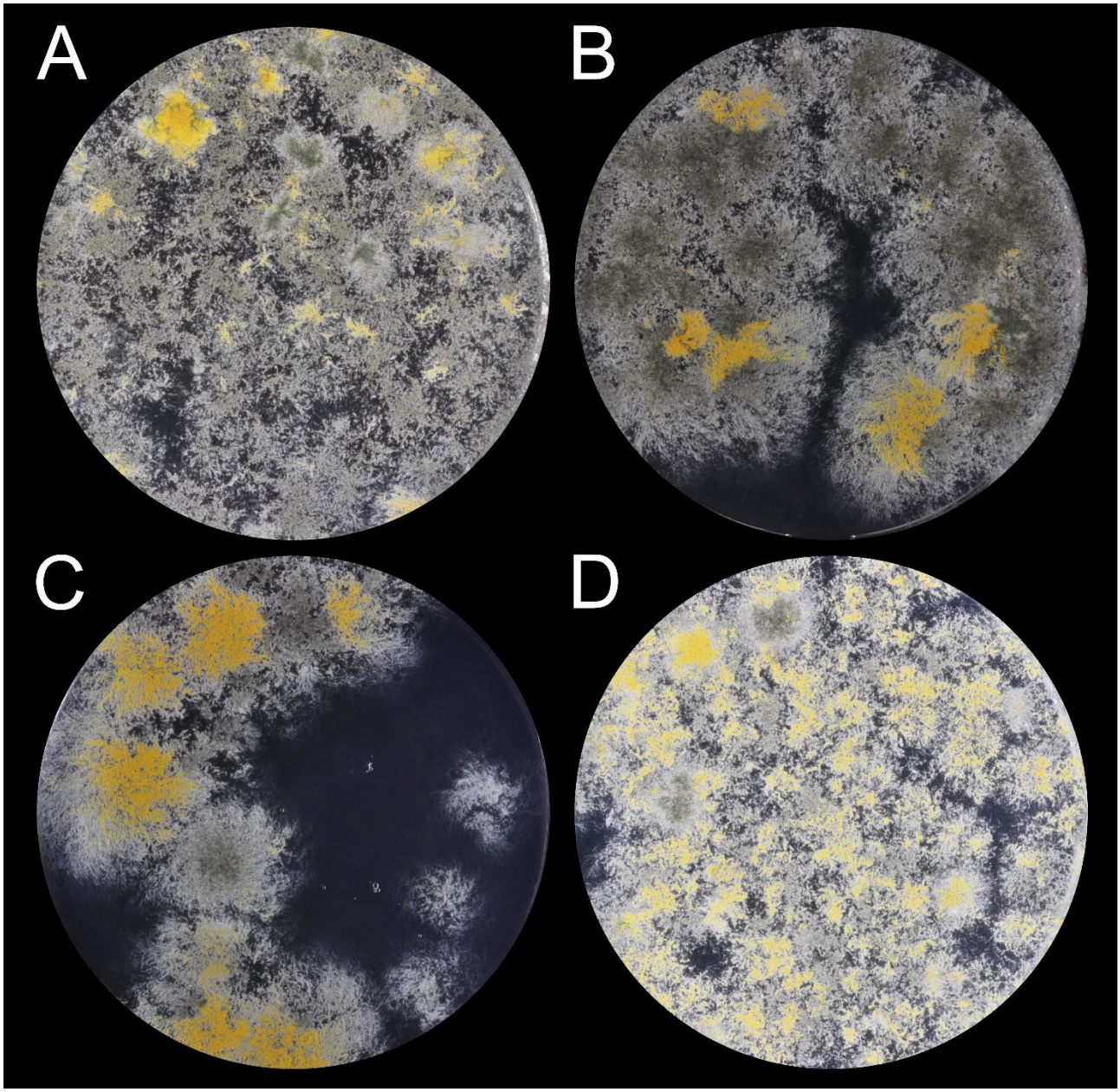

Figure 4.5 Inverting selection by using CRISPR-Cas9. Co-transformation of pFC334 and p6f-yA, selecting for pFC334 on MM-uri-ura plates. A) NID5, linear gene targeting substrate. B) NID5, circular gene targeting substrate. C) NID1, linear gene targeting substrate. D) NID1, circular gene targeting substrate.

Using CRISPR-cas9 to introduce seamless point mutations.

Another application which would be extremely useful, is if CRISPR-Cas9 can be used to introduce point mutations, here very loosely defined as small localized changes in a small numbers of bases, since compared to making insertions or deletion, making point mutations is a bit more tricky. Especially if no counter-selectable marker is available, it is often necessary to leave behind a

selection marker close to the gene being point mutated, which can possible affect expression levels by proximal effects. In any case, introducing point mutations tend to be more time consuming than insertions or deletions, typically requiring either additional cloning steps or transformation steps. An interesting aspect is that since the desired change is so small, it is possible that a ssDNA oligo could potentially be used as repair template, greatly speeding up the process, saving time spent constructing traditional gene targeting vectors. This strategy has been applied to genome engineering both with zinc-finger nucleases and CRISPR-Cas9 [29,30].

One thing to consider is that while a gene can be disrupted in many different ways, repairing a disabled gene requires much more precision. To be usable, for instance for protein engineering, precision is crucial. In chapter three, two of the strains created by mutagenesis by NHEJ, NID1829 and NID1830 contain the same 2 bp deletion in *yA*, causing a frameshift leading to the yellow-spored phenotype. To get an assay, which can be used to accurately, measure correct repair, it was decided to use this specific mutation and try to repair this 2 bp deletion, since frameshifts are much harder to spontaneously revert than base substitutions.

A protospacer was designed to target the mutated allele but not the wildtype of *yA*, and the expression cassette cloned into pFC330, giving the vector pCRISPR1-13p. Two oligos of 60 nucleotide (nt) (CSN520) and 90 nt (CSN518) were designed to correspond to the wildtype *yA* allele, flanking 30 and 45 nt on each side of the mutation site. Furthermore, three vectors were created containing parts of the wildtype *yA*, with flanks, centered from the mutation site of 500 bp, 1000 bp and 2000 bp respectively, named p1-yA500, p1-yA1000 and p1-yA2000 respectively. One thing to notice is that the vector p1-yA2000, due to the relatively short length of *yA*, 2147 bp including introns, as a consequence contains the entire gene and approximately 1300bp up and approximately 500 bp downstream sequence, as such it is possible it is still functional if integrated ectopically.

The strain NID1829 was protoplasted and co-transformed with pCRISPR1-13p and either one of the two ssDNA oligos or one of the three vectors, all designed to provide a template for repair of the 2 bp deletion in *yA*. The transformations plate can be seen in Figure 4.6. Panel A and B shows the use of oligo templates, CSN520 and CSN518 of 60 and 90 bp length respectively. Panel C, D and E shows circular p1-yA500, p1-yA1000 or p1-yA2000 used for repair template.

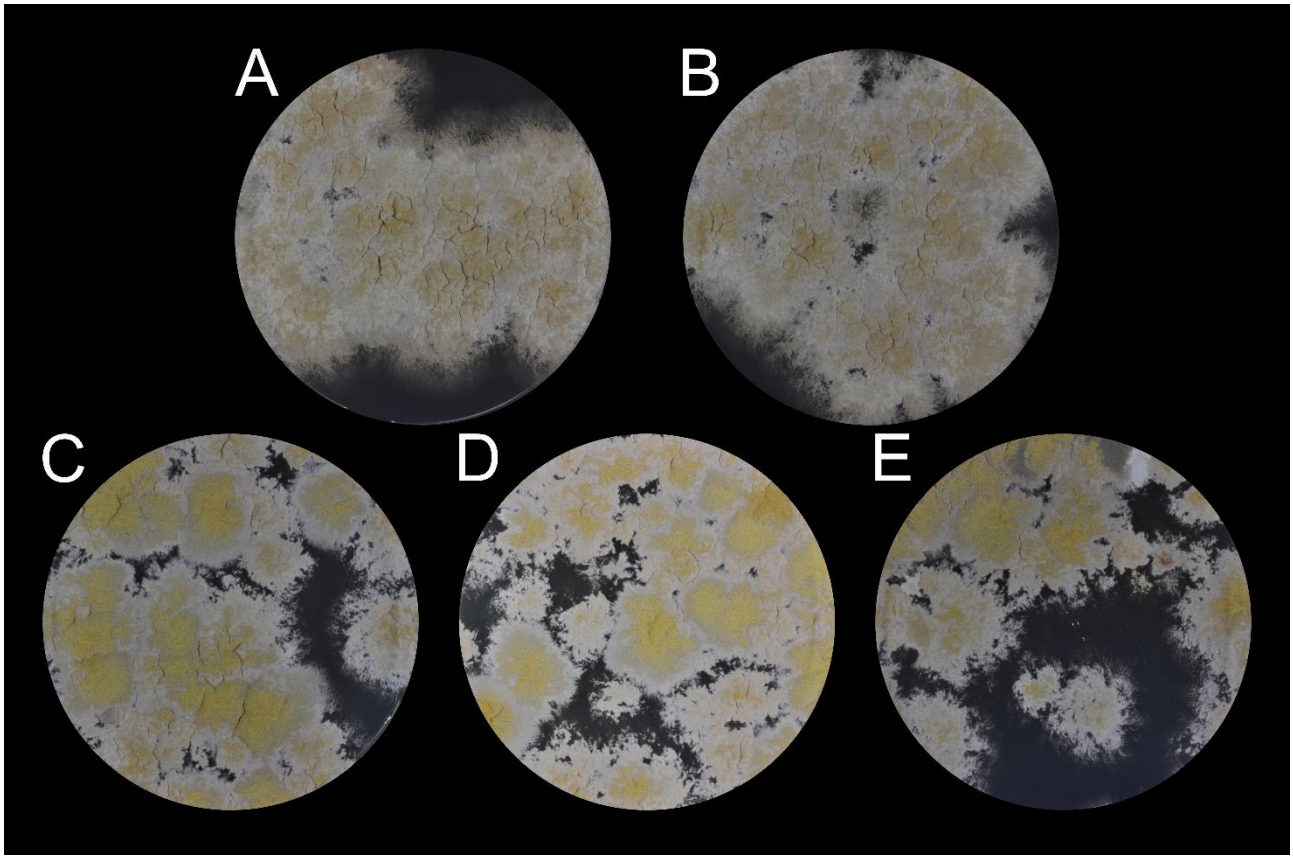

Figure 4.6 Repair of *yA* in an NHEJ proficient background. NID1829 transformed with pCRISPR1-13p, targeting the yA mutant allele and different gene targeting substrate for repair of *yA*. A) and B) are using ss oligos of either 60 or 90 nucleotides in length respectively. C), D) and E) are using circular vector encoding parts of the *yA* gene, with flanks from each side of the mutation being either 500, 1000 or 2000 bp respectively.

A single green colony from the transformation with the 90 nt oligo, showed that using oligos as a repair template for HR is possible, although the rate of correct repair is very low. Similarly, the circular substrate p1-yA2000 gave rise to a few green colonies, however the conclusion is less clear, since ectopic integration, possible in a NHEJ proficient background, can lead to heterologous expression due the long flank length. However, there is no selection pressure to integrate the fragment.

Since it showed that the oligo was a valid repair template, two 90 nt ssDNA oligos was designed, to either introduce a stop codon (CSN486) or the same 2bp deletion as the one in NID1829 (CSN519). NID1 was co-transformed with pFC334 and either of the two oligos. Transformation plates can be seen in Figure 4.7, panel A showing transformation with an oligo introducing a stopcodon and the

panel B showing transformation with an oligo encoding the 2 nt deletion corresponding to the one in NID1829.

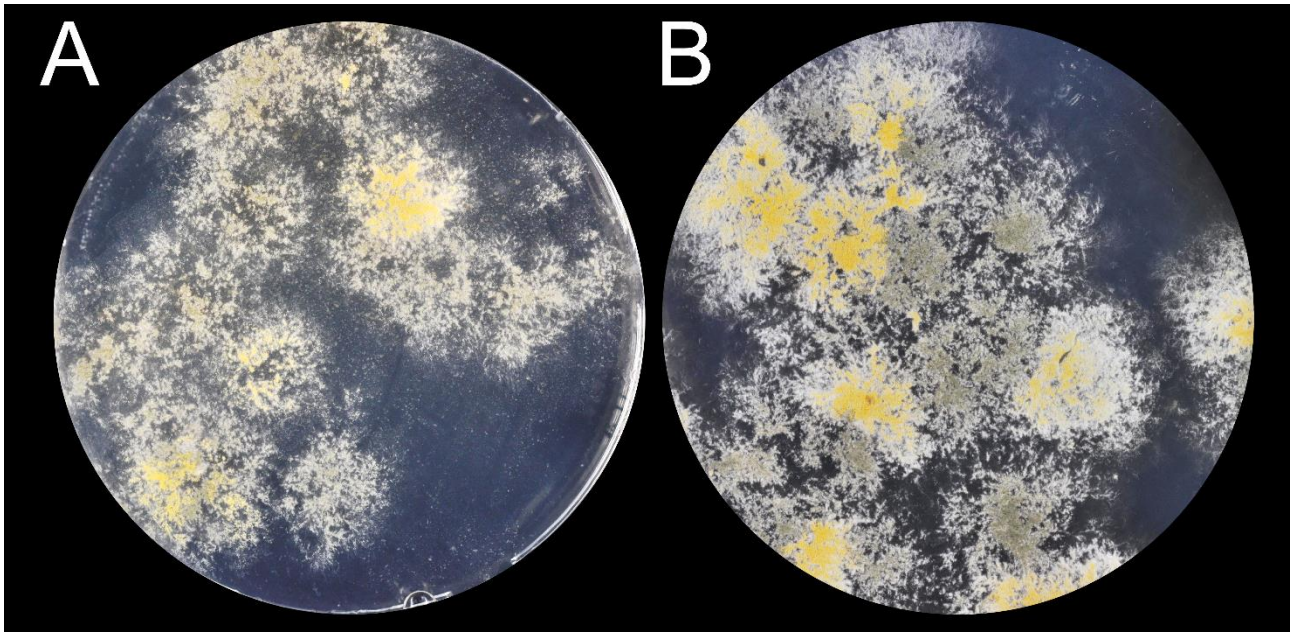

Figure 4.7 Introducing a specific point mutation using oligos in and NHEJ deficient background. NID1 transformed with pFC334 and a 90 nt oligo designed to introduce either A) a stop codon in *yA* or B) the same 2 bp deletion as found in NID1829.

The transformations resulted in approximate 40% yellow colonies, however both green and yellow colonies were even more uneven in appearances than normally seen when transforming with AMA1 based plasmids. Five colonies from the transformation with the 2 bp deletion oligo, CSN519, were sequenced in the region containing the expected mutation. Three of the five had the correct 2 bp deletion as expected from the oligo, while the two others, while mutated in the locus, had a different deletions. This shows, that not surprisingly, this strategy is much more efficient in a NHEJ deficient background, but also that in lieu of mutagenesis by faulty NHEJ, which is not possible in a NHEJ deficient recipient, oligo-directed mutagenesis can be used as a substitute.

Next, one of the strains carrying the 2 bp deletion, NID2028 was protoplasted and co-transformed with pCRISPR1-13p and either of the two oligos, CSN518 or CSN520, or one of the three vectors p1-yA500, p1-yA1000 or p1-yA2000, either linearized or in circular form. The transformation plates can be seen in Figure 4.8. Panel A and B shows CSN520 and CSN518, panel C, D and E shows linearized p1-yA500, p1-yA1000 or p1-yA2000 as template and panel F, G and H shows circular p1-yA500, p1-yA1000 or p1-yA2000 used for repair template.

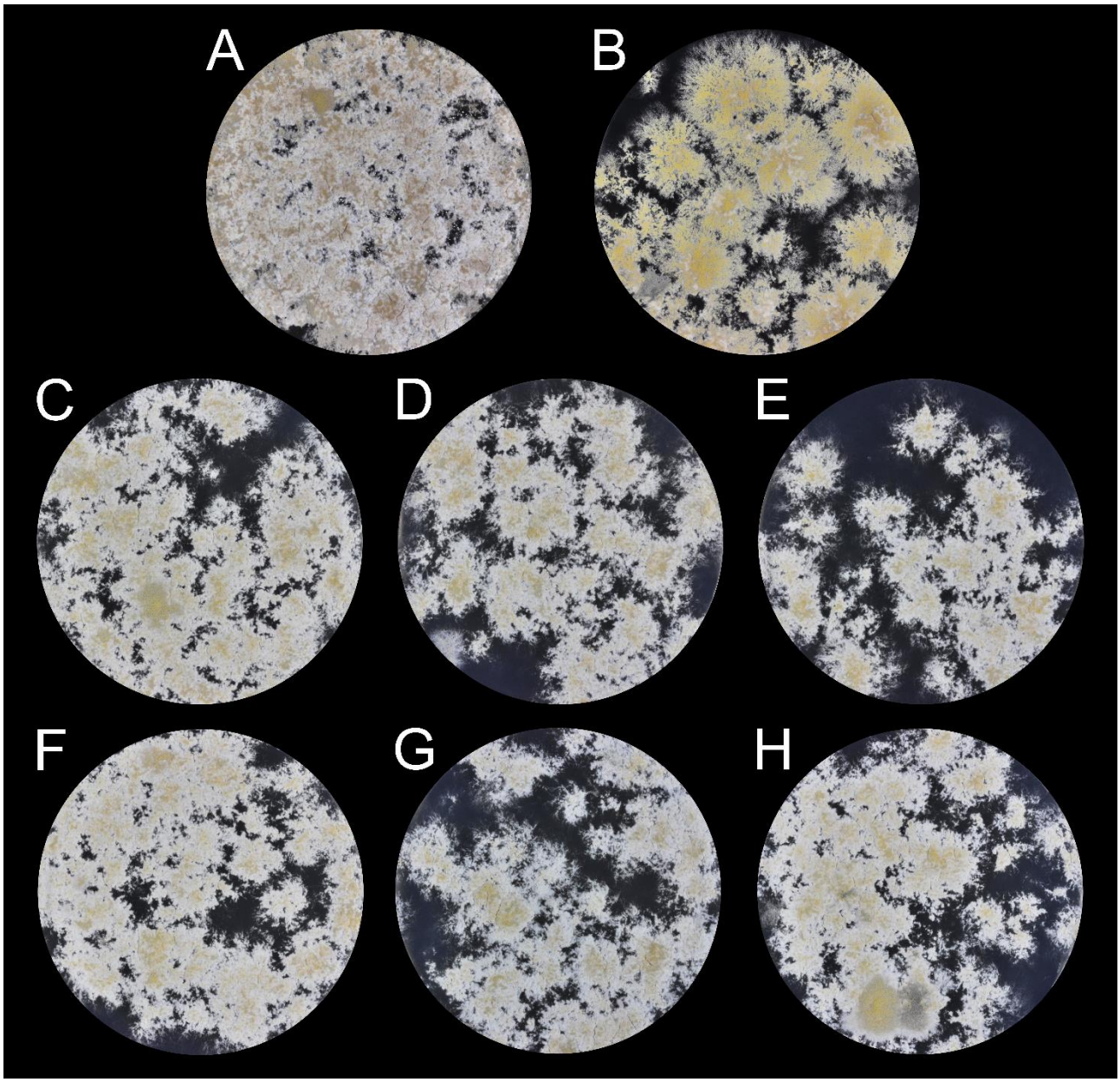

Figure 4.8 Repair of *yA* in a NHEJ deficient background. NID2028, which is NHEJ deficient but otherwise identical to NID1829, transformed with pCRISPR1-13p, targeting the *yA* mutant allele and different gene targeting substrate for repair of *yA*. A) and B) are using ss oligos of either 60 or 90 nucleotides in length respectively. C), D) and E) are using linearized vector encoding parts of the *yA* gene, with flanks from each side of the mutation being either 500, 1000 or 2000 bp respectively. F), G) and H) are using circular vector encoding parts of the yA gene, with flanks from each side of the mutation being either 500, 1000 or 2000 bp respectively.

Considering the relatively high rate of integration of the *yA* deletion vector p6f-yA and also correct introduction of the 2 bp deletion with the CSN519 oligo, the low frequency of repair of *yA* is somewhat surprising.

There are several parameters to consider when designing gRNAs, and a great variety of online tools exist. In organisms with large genomes, e.g. plants and animals, often the main concern is limiting off-targeting by finding the most unique sites for targeting within the gene of interest. Other studies have focused on optimizing on-target efficiency [31,32], finding the parameters leading to the most efficient gRNAs. However, when using CRISPR-Cas9 for introducing point mutations, the choice of gRNA is often quite limited since it needs to target the gene of interest but not the template used for introducing the point mutation. In this case only two protospacers are available without introducing silent mutations in the repair template. The other possible protospacer was designed and cloned into pFC330, creating the vector pCRISPR1-20p. The experiments were repeated with the new vector, however the frequency of repair did not change significantly.

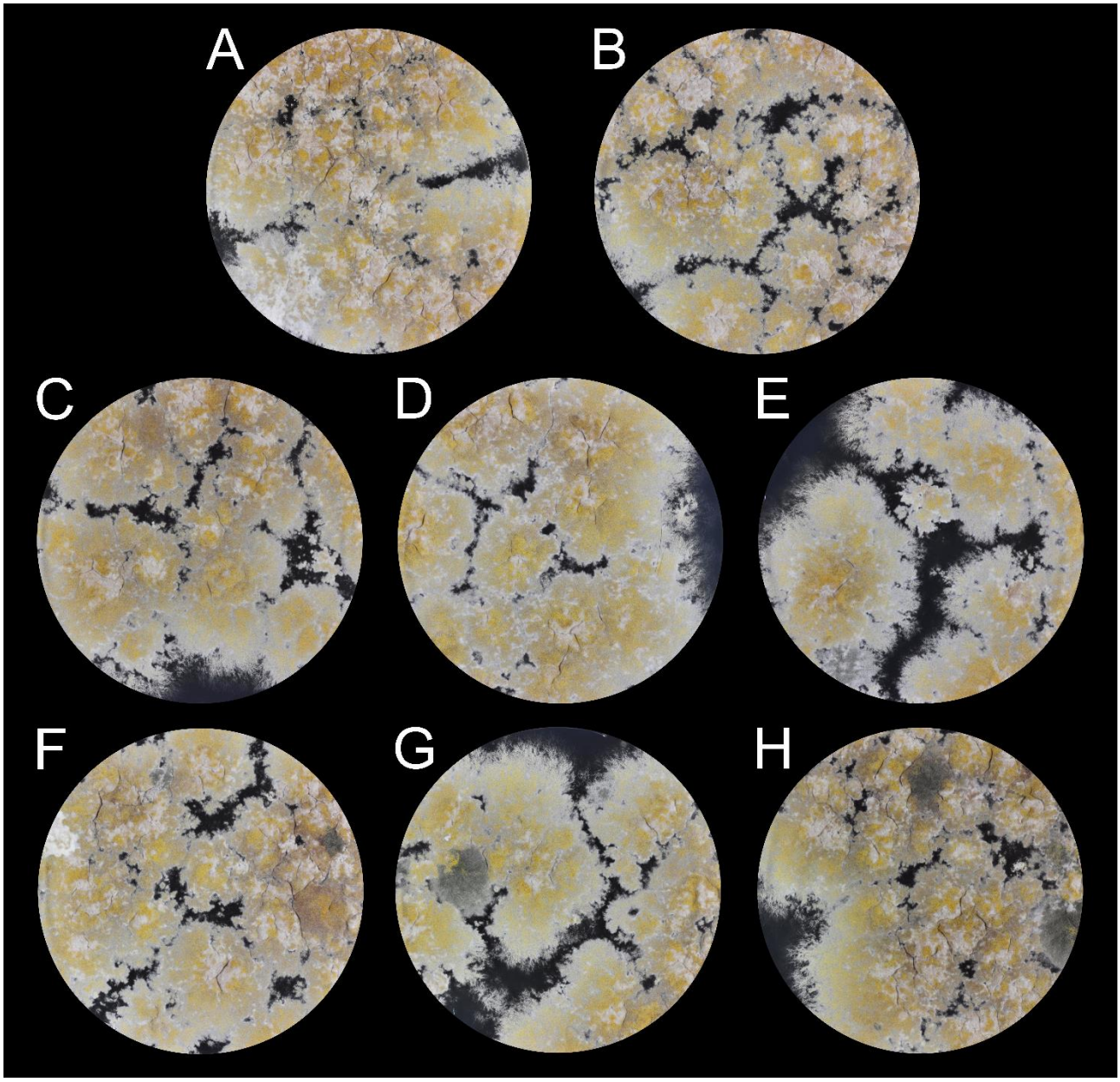

Figure 4.9 Repair of *yA* in a NHEJ deficient background. NID2028 is transformed with pCRISPR1-20p, targeting the same *yA* mutant allele as pCRISPR1-13p, but with a different protospacer and different gene targeting substrate for repair of *yA*. A) and B) are using ss oligos of either 60 or 90 nucleotides in length respectively. C), D) and E) are using linearized vector encoding parts of the yA gene, with flanks from each side of the mutation being either 500, 1000 or 2000 bp respectively. F), G) and H) are using circular vector encoding parts of the *yA* gene, with flanks from each side of the mutation being either 500, 1000 or 2000 bp respectively.

Overall, these results show that introducing small changes such deletions, insertions or substitutions are possible with both long ssDNA oligos (90nt), but also linearized or circular vectors. However due to the low frequency in the *yA* repair assay and the high colony density on the plates, it is not possible to say anything specific on the frequencies. However the tendencies seem to be similar to what is observed when using the *yA* deletion vector, that a circular substrate works better than linear

substrate[3], and also longer flanks seem to be more efficient than shorter flanks [22]. It seems that repairing *yA* was much more inefficient compared to mutating it, even considering that not every yellow colony had the specific desired mutation, although again there is not enough data for anything conclusive. However there is no doubt that it is going to be an extremely useful technique, especially if further studies reaffirms that ssDNA oligos can be used with high gene targeting frequencies, greatly speeding up the process.

Discussion

Currently six papers on using CRISPR-Cas9 for genome editing in filamentous fungi, including the one presented in chapter 3, have been published. What is really interesting is that the main challenge, introducing the components of the system, particularly the gRNA, was solved in widely different manners. In *Trichoderma reesei*, in a work by Liu and co-workers [1], a codon-optimized Cas9 was randomly integrated using either a constitutive or an inducible promoter and the gRNA made by in vitro transcription and introduced into protoplasts. In *Pyricularia oryzae* (*Magnaporthe oryzae*), in a study by Arazoe and co-workers [2] again a codon-optimized Cas9 was used, but this time inserted in a non-integrative, transiently expressed vector together with the constructs necessary for expression of the gRNA which is done either from a *P. oryzae* U6-1 or U6-2 RNA polymerase III promoter or the RNA polymerase II promoter *trpC*. In *Neurospora crassa* by Matsu-ura and co-workers [4], transient plasmid expression was similarly used, however for expression of the gRNA, the *S. cerevisiae* SNR52 promoter was used. In *Aspergillus fumigatus* by Fuller and coworkers [33] the humanized Cas9, under control of the S. cerevisiae *TEF1* promoter and the gRNA under control of the *S. cerevisiae SNR52* promoter [34] was integrated into the genome by random integration, in either one or two steps. In *Ustilago maydis* by Schuster and co-worker [35] expression from transient plasmids was also used, and an *U. maydis* U6 promoter was used for sgRNA expression. Comparing these systems show huge differences in what works and what does not work. In *T. reesei* codon-optimization was found to be necessary [1], while the humanized Cas9 worked in *A. fumigatus* [33]. Likewise, while strategies for expressing the gRNA, including expressing from RNA polymerase II promoters directly, fungal U6 promoters or the *S. cerevisiae SNR52* promoter proved unsuccessful in *A. nidulans* as described in this study, while successful in

other organisms. Overall this leaves room for lots of optimizations, figuring out what works and what does not work, and why.

Another interesting concern is the delivery of Cas9 and the sgRNA. Of the two components, Cas9 and the sgRNA, Cas9 is the same every time, while the sgRNA changes with every target, greatly affecting the strategies used for expressing them. For expression of the Cas9 protein, that has been used in filamentous fungi so far, is either genomic integration or transient expression from plasmids. Considering stability, genomic integration is by far the most stable way of expression, since once a pure strain is obtained, then one copy of the Cas9 gene will be present in every single nuclei. The disadvantage however that it will be constitutively present, and while no negative impact as a result of Cas9 expression, has been observed in filamentous fungi so far, in *S. cerevisiae* reduced growth rates have been observed as a result of strong Cas9 expression [36]. A way to circumvent this is by using inducible promoters for Cas9 expression, such as done in *T. reesei* [1]. Alternatively, with transient expression, which so far in all cases have been DNA based, the DNA carrying the various elements is lost again, leaving a cleaner strain. However expression is probably uneven and there is the risk that the DNA will integrate into the genome. Along the same lines of thought, so far all the systems have been dependent on selection, since when doing transformation only a subset of cells are actually competent. That leaves three options, a selectable phenotype, the DNA vehicle for expression one or both of the component or the gene targeting template in the cases of HR. The most flexible of these is the DNA vehicle, but again it might increase the risk of genomic integration. And again it shows that there are lots of parameter to consider.

For the purpose of making insertions, deletions and point mutations without the integration of a selection marker, several types of gene targeting substrates have been tested. If looking purely at optimizing the frequency of correct integrations, the tendency was that circular substrates with long flanks did seem to be the most efficient. However when designing an optimal strategy for genetic engineering, the gene targeting frequency is only one parameter. Different gene targeting substrates have different advantages. While the most efficient, circular gene targeting substrates tends to be the most time consuming to make, requiring a cloning step. However if the gene targeting substrate is complex, consisting of many different parts, it might be the only way to go. Linear substrates can, if simple, be made by fusion PCR, speeding up the process, while more complex ones are easiest made by cloning a vector and subsequent linearization, e.g. by restriction enzymes. ssDNA oligos can be made synthetically, and are commercially available. While limited by their small size, they

can easily incorporate a point mutation or a short tag or even more interestingly is the option of incorporation of randomized nucleotides, for instance to vary a single amino acid at a specific position in a protein. However the tradeoff seem to be lower gene targeting efficiency, compared to a longer double-stranded gene targeting substrate. Even greater effect is the presence or absence of NHEJ. While no attempts on multiplexing have been done during these studies, the foundation have been laid with the exploration of gene targeting without an accompanying selection marker. However, it is reasonable to assume that the differences in gene targeting efficiencies will only be exacerbated, as the number of targets go up. Therefore, one of the most important parameters when deciding on a genetic engineering strategy is the screening system available. With high-throughput methods available or with a phenotype to screen for, compromises to gene targeting efficiency can be made. However the less efficient the screening method is, the more crucial it is to have high gene targeting frequencies, so that correct transformants can be identified in a timely manner.

Another aspect of interest is off-targeting, especially since the verdict is still out to whether it is a problem or not. Here two methods for limiting off-targeting were tested. One was based on shortening the protospacer from 20 bp down to 17, which has been shown to be a viable strategy in human cells [10]. However this lead to aberrant, sporeless colonies in A. nidulans. While protospacers of 18 or 19 bases have not been tested, using 17 bp is certainly not an effective strategy, despite it being an otherwise simple and elegant solution if it did work as in human cells. The other method is based on creating two nicks in close proximity, rather than a single DSB. When used to facilitate integration of a gene targeting substrate by HR, two subpopulations of colonies arose. One was tiny and sporeless, while the other appeared normal in morphology and showed a similar yellow/green distribution as that observed when stimulating gene targeting with wildtype Cas9. While the small colonies is a bit worrisome, further analysis should be able to shed light upon whether any negative effects is confined to those, in which case it is not a problem since they are easily separated, or if the normal colonies also are somehow affected. If further analysis reveals no issues with normal sized colonies, it is good a strategy to minimize off-targeting, when doing gene targeting. It has two disadvantages compared to wildtype Cas9, however. The nickase variant, unlike the wildtype does not seem to induce mutagenesis by NHEJ if no template is available, somewhat limiting the applications, although not greatly so, since applications available for gene targeting by HR are more versatile. The other disadvantage is that because the need for expressing two sgRNAs rather than one, the cost and complexity of creating the vectors goes up. However

further studies will have to be done to figure out the degree of which off-targeting is a concern in Aspergillus species.

Overall this study explores some more aspects of using CRISPR-Cas9 for genetic engineering in *Aspergillus*, and especially the prospect of not having to integrate a selection marker as a part of the gene targeting substrate and be able to use ssDNA oligos as a repair template is exciting.

Materials and methods

Strains and media

Escherichia coli strain DH5α was used to propagate all plasmids. The *Aspergillus nidulans* strains used in this study are listed in Table 1. Genomic DNA (gDNA) from fungal strains was isolated via FastDNA SPIN Kit for Soil DNA extraction kit (MP Biomedicals, USA). All strains were cultivated on standard glucose based minimal medium (MM) (1% glucose, 1x nitrate salt solution [37], 0.001% Thiamine, 1x trace metal solution [38], 2% agar), supplemented with 10mM uridine (Uri), 10mM uracil (Ura), and/or 4mM L-arginine (Arg) when required. Solid plates containing 5-fluoroorotic acid (5-FOA) were made as MM+Arg+Uri+Ura supplemented with filter-sterilized 5-FOA (Sigma-Aldrich) to a final concentration of 1.3 mg/ml. For transformation media (TM) glucose was replaced with 1M sucrose.

Table 4.2 Strains used in the study

NID1	*argB2, veA1, pyrG89, nkuAΔ*
NID5	*argB2, veA1, pyrG89*
NID1829	*argB2, veA1, pyrG89, yA**
NID2028	*argB2, veA1, pyrG89, nkuAΔ, yA**

PCR and USER cloning

All vectors were constructed by PCR as described by Nødvig et al [3] and USER cloning as described by Hansen et al [39]. Detailed vector construction can be found in appendix 1 and the primers used in appendix 2. An overview of the vectors can be found in Table 3.

Table 4.3 vectors used in the study

Name	Genotype	Source
pFC330	*AMA1-pyrG-Cas9*-PacI/Nt.BbvCI cassette	[3]
pFC331	*AMA1-argB-Cas9* PacI/Nt.BbvCI cassette	[3]
pFC334	*AMA1,argB, Cas9, PS1*	[3]
p6f-yA	*yA-up, Afl-pyrG, yA-dw*	[3]
pCRISPR-gpdA-1a	*AMA1,argB, Cas9, PS1 (gpdA)*	This study
pCRISPR-SNR52-1a	*AMA1,argB, Cas9, PS1 (SNR52)*	This study
pCRISPR-U6-1-1a	*AMA1,argB, Cas9, PS1 (U6-1)*	This study
pCRISPR-U3-1a	*AMA1,argB, Cas9, PS1 (U3)*	This study
pCRISPR1-13p	*AMA1, pyrG, Cas9, PS13*	This study
pCRISPR1-20p	*AMA1, pyrG, Cas9, PS20*	This study
pCRISPR1-14p	*AMA1, pyrG, Cas9, PS14*	This study
pCas9-D10A-argB	*AMA1- argB -Cas9(D10A)*-PacI/Nt.BbvCI cassette	This study
pCas9-H840A-argB	*AMA1- argB -Cas9(H840A)*-PacI/Nt.BbvCI cassette	This study
pCas9-D10A-pyrG	*AMA1-pyrG-Cas9(D10A)*-PacI/Nt.BbvCI cassette	This study

pCas9-H840A-pyrG	*AMA1-pyrG-Cas9(H840A)*-PacI/Nt.BbvCI cassette	This study
pCRISPRn1-10-11a	*AMA1- argB -Cas9(D10A), PS10-PS11*	This study
pCRISPRn1-11-12a	*AMA1- argB -Cas9(D10A), PS11-PS12*	This study
pCRISPRn2-10-11a	*AMA1- argB -Cas9(H840A), PS10-PS11*	This study
pCRISPRn2-11-12a	*AMA1- argB -Cas9(H840A), PS11-PS12*	This study

Table 4.4 Protospacers used in the study

Protospacer	**Sequence**	**Gene**
PS1	GGCGGAGTATCATAACATCG	*yA*
PS10	ATCAATAAGACTGCCTCAAT	*yA*
PS11	GCGACCCTACGCGGTTATAA	*yA*
PS12	GTATTTCTGCTAACTCTCTC	*yA*
PS13	TCGGCGGAGTATCATAATCG	*yA* (mutated)
PS14	GGAGTATCATAACATCG	*yA*
PS20	ATCATAATCGAGGTTGAGTC	*yA* (mutated)

Transformation and strain validation by Tissue-PCR

Protoplastation was performed as described by Nielsen et al [23]. Transformation using either *pyrG* or *argB* as genetic marker was performed as described by Nødvig et al [3]. One note, in the case of inverted selection, approximately 1 µg of CRISPR-Cas9 vector and approximately 3 µg gene targeting substrate were used.

Acknowledgement

I thank my bachelor student Charlotte Beck for the creation of the eight plasmids used for testing Cas9 nickase variant.

References

1. Liu R, Chen L, Jiang Y, Zhou Z, Zou G. Efficient genome editing in filamentous fungus Trichoderma reesei using the CRISPR/Cas9 system. Cell Discov. 2015;1: 15007. doi:10.1038/celldisc.2015.7

2. Arazoe T, Miyoshi K, Yamato T, Ogawa T, Ohsato S, Arie T, et al. Tailor-made CRISPR/Cas system for highly efficient targeted gene replacement in the rice blast fungus. Biotechnol Bioeng. 2015; n/a–n/a. doi:10.1002/bit.25662

3. Nødvig CS, Nielsen JB, Kogle ME, Mortensen UH. A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi. Yu J-H, editor. PLoS One. 2015;10: e0133085. doi:10.1371/journal.pone.0133085

4. Matsu-ura T, Baek M, Kwon J, Hong C. Efficient gene editing in Neurospora crassa with CRISPR technology. Fungal Biol Biotechnol. BioMed Central; 2015;2: 4. doi:10.1186/s40694-015-0015-1

5. Ran FA, Hsu PD, Lin C-Y, Gootenberg JS, Konermann S, Trevino AE, et al. Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. Cell. Elsevier; 2013;154: 1380–1389. doi:10.1016/j.cell.2013.08.021

6. Shen B, Zhang W, Zhang J, Zhou J, Wang J, Chen L, et al. Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. Nat Methods. 2014;11: 399–402. doi:10.1038/nmeth.2857

7. Tsai SQ, Wyvekens N, Khayter C, Foden J a, Thapar V, Reyon D, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nat Biotechnol. 2014;32: 569–576. doi:10.1038/nbt.2908

8. Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. Nat Biotechnol. 2014;32: 577–582. doi:10.1038/nbt.2909

9. Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol. 2013;31: 833–838. doi:10.1038/nbt.2675

10. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using

truncated guide RNAs. Nat Biotechnol. 2014;32: 279–284. doi:10.1038/nbt.2808

11. Cradick TJ, Fine EJ, Antico CJ, Bao G. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. Nucleic Acids Res. 2013;41: 9584–9592. doi:10.1093/nar/gkt714

12. Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna J a, Liu DR. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat Biotechnol. Nature Publishing Group; 2013;31: 839–43. doi:10.1038/nbt.2673

13. Jakočiūnas T, Bonde I, Herrgård M, Harrison SJ, Kristensen M, Pedersen LE, et al. Multiplex metabolic pathway engineering using CRISPR/Cas9 in Saccharomyces cerevisiae. Metab Eng. 2015;28: 213–222. doi:10.1016/j.ymben.2015.01.008

14. Grav LM, Lee JS, Gerling S, Kallehauge TB, Hansen AH, Kol S, et al. One-step generation of triple knockout CHO cell lines using CRISPR/Cas9 and fluorescent enrichment. Biotechnol J. 2015;10: 1446–1456. doi:10.1002/biot.201500027

15. Ronda C, Maury J, Jakočiūnas T, Baallal Jacobsen SA, Germann SM, Harrison SJ, et al. CrEdit: CRISPR mediated multi-loci gene integration in Saccharomyces cerevisiae. Microb Cell Fact. 2015;14: 97. doi:10.1186/s12934-015-0288-3

16. Mans R, van Rossum HM, Wijsman M, Backx A, Kuijpers NG a., van den Broek M, et al. CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in Saccharomyces cerevisiae. FEMS Yeast Res. 2015;15: fov004–fov004. doi:10.1093/femsyr/fov004

17. Ballance DJ, Buxton FP, Turner G. Transformation of Aspergillus nidulans by the orotidine-5'-phosphate decarboxylase gene of Neurospora crassa. Biochem Biophys Res Commun. 1983;112: 284–289. doi:10.1016/0006-291X(83)91828-4

18. Boeke JD, La Croute F, Fink GR. A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. MGG Mol Gen Genet. 1984;197: 345–346. doi:10.1007/BF00330984

19. Storici F, Resnick M a. The Delitto Perfetto Approach to In Vivo Site-Directed Mutagenesis and Chromosome Rearrangements with Synthetic Oligonucleotides in Yeast. Methods in Enzymology. 2006. pp. 329–345. doi:10.1016/S0076-6879(05)09019-1

20. Horecka J, Davis RW. The 50:50 method for PCR-based seamless genome editing in yeast. Yeast. 2013: 103–112. doi:10.1002/yea.2992

21. Hua SB, Qiu M, Chan E, Zhu L, Luo Y. Minimum length of sequence homology required for in vivo cloning by homologous recombination in yeast. Plasmid. 1997;38: 91–96. doi:10.1006/plas.1997.1305

22. Krappmann S. Gene targeting in filamentous fungi: the benefits of impaired repair. Fungal Biol Rev. 2007;21: 25–29. doi:10.1016/j.fbr.2007.02.004

23. Nielsen ML, Albertsen L, Lettier G, Nielsen JB, Mortensen UH. Efficient PCR-based gene targeting with a recyclable marker for Aspergillus nidulans. Fungal Genet Biol. 2006;43: 54–64. doi:10.1016/j.fgb.2005.09.005

24. Weyda I, Lübeck M, Ahring BK, Lübeck PS. Point mutation of the xylose reductase (XR) gene reduces xylitol accumulation and increases citric acid production in Aspergillus carbonarius. J Ind Microbiol Biotechnol. 2014;41: 733–739. doi:10.1007/s10295-014-1415-6

25. Arnaud MB, Cerqueira GC, Inglis DO, Skrzypek MS, Binkley J, Chibucos MC, et al. The Aspergillus Genome Database (AspGD): Recent developments in comprehensive multispecies curation, comparative genomics and community resources. Nucleic Acids Res. 2012;40: D653–D659. doi:10.1093/nar/gkr875

26. Belhaj K, Chaparro-Garcia A, Kamoun S, Nekrasov V. Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. Plant Methods. Plant Methods; 2013;9: 39. doi:10.1186/1746-4811-9-39

27. Brow D a, Guthrie C. Transcription of a yeast U6 snRNA gene requires a polymerase III promoter element in a novel position. Genes Dev. 1990;4: 1345–1356. doi:10.1101/gad.4.8.1345

28. Fauser F, Schiml S, Puchta H. Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. Plant J. 2014;79: 348–359. doi:10.1111/tpj.12554

29. Chen F, Pruett-Miller SM, Huang Y, Gjoka M, Duda K, Taunton J, et al. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. Nat Methods. 2011;8: 753–755. doi:10.1038/nmeth.1653

30. Ran FA, Hsu PD, Wright J, Agarwala V, Scott D a, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8: 2281–2308. doi:10.1038/nprot.2013.143

31. Gagnon J a., Valen E, Thyme SB, Huang P, Ahkmetova L, Pauli A, et al. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. PLoS One. 2014;9: 5–12. doi:10.1371/journal.pone.0098186

32. Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol. Nature Publishing Group; 2014;32. doi:10.1038/nbt.3026

33. Fuller KK, Chen S, Loros JJ, Dunlap JC. Development of the CRISPR/Cas9 system for targeted gene

disruption in Aspergillus fumigatus. Eukaryot Cell. 2015; EC.00107–15. doi:10.1128/EC.00107-15

34. Dicarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. 2013;41: 4336–4343. doi:10.1093/nar/gkt135

35. Schuster M, Schweizer G, Reissmann S, Kahmann R. Genome editing in Ustilago maydis using the CRISPR–Cas system. Fungal Genet Biol. Elsevier Inc.; 2015; 1–7. doi:10.1016/j.fgb.2015.09.001

36. Ryan OW, Skerker JM, Maurer MJ, Li X, Tsai JC, Poddar S, et al. Selection of chromosomal DNA libraries using a multiplex CRISPR system. Elife. 2014;3: e03703. doi:10.7554/eLife.03703

37. Kaminskyj SGW. Fundamentals of growoth, storage, genetics and microscopy of Aspergillus nidulans. Fungal Genet Newsl. 2001; 25–31.

38. Cove DJ. The induction and repression of nitrate reductase in the fungus Aspergillus nidulans. Biochim Biophys Acta. Elsevier Publishing Company, Amsterdam; 1966;113: 51–56. doi:10.1016/S0926-6593(66)80120-0

39. Hansen BG, Salomonsen B, Nielsen MT, Nielsen JB, Hansen NB, Nielsen KF, et al. Versatile enzyme expression and characterization system for Aspergillus nidulans, with the Penicillium brevicompactum polyketide synthase gene from the mycophenolic acid gene cluster as a test case. Appl Environ Microbiol. 2011;77: 3044–3051. doi:10.1128/AEM.01768-10

Appendix 4.1 Vector construction

Vectors for testing different promoter for sgRNA expression

A gBlock encoding the same *yA* specific sgRNA as pFC334, protospacer 1, with the *gpdA* promoter and *trpC* terminator was ordered from IDT (IDTdna, Belgium), PCR amplified with the primers CSN389+CSN390 and cloned into pFC331 yielding the vector pCRISPR1-gpdA-1a.

Another gBlock encoding the same *yA* specific sgRNA but this time with the S. cerevisiae SNR52 promoter and SUP4 terminator was likewise amplified with primer CSN389+CSN390, cloned into pFC331, yielding the vector pCRISPR-SNR52-1a.

A third gBlock with the same *yA* specific sgRNA, but flanked with 307 bp upstream and 53 bp downstream of the A. fumigatus U6-1 snRNA, to constitute the U6-1 promoter and terminator, was amplified with primer CSN389+CSN390 and cloned into pFC331, yielding the vector pCRISPR-U6-1-1a.

The vector pCRISPR-U3-1a was made by amplifying the 486 bp upstream of the U3 snoRNA of *A. fumigatus* from *A. fumigatus* gDNA with primers CSN438+CSN439, while the yA sgRNA and the SUP4 terminator was amplified with primers CSN363+CSN390 using pCRISPR-SNR52-1a as template. Fragments were cloned into pFC331 by USER cloning, giving the vector pCRISPR-U3-1a

Cas9 nickase vectors

Cas9 including the A. nidulans tef1 promoter and terminator was amplified in two part introducing either the D10A or the H840 mutation with the primer tails. For the D10A mutation upstream part of Cas9 was amplified with primers CSN385+CSN351 and downstream with primers CSN352+CSN323. Similarily the H840A was introduced by amplifying 5’ end with primer CSN385+CSN353 and 3’end with primers CSN354+CSN323, all four fragments using pU0000-Cas9 as template [1]. Additionally a fragment containing either the *argB* or *pyrG* fungal selection marker and part of the ampicillin bacterial selection marker was amplified from CSN381+CSN380 or CSN382+CSN380 with pAC76 or pAC161 (Aspergillus plasmid collection) and two fragments containing the remaining part of the ampicillin marker gene and the 5’ end of the AMA1 element

and one containing the 3' end of AMA1 was amplified with primer CSN379+KBR087 and KBR088+CSN386 using pAC76 as template. From this four vectors were assembled by USER fusion, pCas9-D10A-argB, pCas9-H840A-argB pCas9-D10A-pyrG, pCas9-H840A-pyrG, each containing either *argB* or *pyrG* and either the D10A or H840A mutation.

Three protospacers for targeting *yA*, protospacer 10, 11 and 12 were designed, in such way that the combination PS10+11 yields two nicks 16 nucleotides apart on opposites strands, while PS11+12 yields two nicks on opposite strand 65 nucleotides apart. For expression of the sgRNAs the *A. nidulans* Histone 3 and 4 bidirectional promoter was used. The promoter fragment for the PS10+11 pair was amplified with primers CSN497+CSN498, while the promoter for the PS11+12 pair was amplified with primer CNS499+CSN498. The *trpC* terminator with the gRNA backbone and a HDV ribozyme is amplified using pFC334 as template with primers CSN503+CSN500 for PS10, CSN501+CSN504 for PS11 and CSN502+CSN503. These fragments was assembled in a total of four vectors with a promoter and two terminator fragments in each, in such way that all the vectors have either the PS10+11 or the PS11+12 combination of gRNAs and either the D10A or H840A mutation and all of them the *argB* selection marker, yielding the vectors pCRISPRn1-10-11a, pCRISPRn1-11-12a, pCRISPRn2-10-11a and pCRISPRn2-11-12a.

CRISPR-Cas9 vectors pFC330, pFC331, pFC334 and *yA* deletion vector p6f-yA were constructed as described in chapter three [1]

Vector with alternative sgRNAs, pCRISPR1-13p, pCRISPR1-14p and pCRISPR1-20p

For each vector two PCR fragments were made using pFC334 as template and the fragments were cloned into vector pFC330. Vector pCRISPR1-13p was made with primers CSN389+CSN521 and CSN522+CSN390. Similarly vector pCRISPR1-14p was made with primers CSN389+CSN536 and CSN537+CSN390. And vector pCRISPR1-20p was made with primers CSN389+CSN574 and CSN575+CSN390.

Single stranded oligos for repair templates are primers CSN518 and CSN520 . The double stranded vector was made by amplifying parts of *yA* from *A. nidulans* genomic DNA, using primer CSN568+569 for the 500 bp flanks, CSN570+571 for 1000 bp flanks and CSN572+573 for 2000 bp flanks. These three fragments were separately cloned into vector pU0002[2] by USER cloning.

References

1. Nødvig CS, Nielsen JB, Kogle ME, Mortensen UH. A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi. Yu J-H, editor. PLoS One. 2015;10: e0133085. doi:10.1371/journal.pone.0133085

2. Hansen BG, Salomonsen B, Nielsen MT, Nielsen JB, Hansen NB, Nielsen KF, et al. Versatile enzyme expression and characterization system for Aspergillus nidulans, with the Penicillium brevicompactum polyketide synthase gene from the mycophenolic acid gene cluster as a test case. Appl Environ Microbiol. 2011;77: 3044–3051. doi:10.1128/AEM.01768-10

Appendix 4.2 Primer table

Promoter comparison		
CSN389	PgpdA-pac-up-fwd	GGGTTTAAU GCGTAAGCTCCCTAATTGGC
CSN390	TtrpC-short-pac-dw-rv	GGTCTTAAU GAGCCAAGAGCGGATTCCTC
CSN363	yA-crispr-planD-fwd	AGTATCAU AACATCG GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG
CSN438	Afum-U3p-fwd	GGGTTTAAU GATCACATAGATGCTCGGTTGACA
CSN439	Afum-U3p-PS1-rv	aTGATACUCCGCC GCTGTTAGCAAAATGTTCTGACTTG
Nickase Cas9 vectors		
CSN323	Anid-Ttef1-L3-rv	ATTACACTU GTATTGGGATGAATTTTGTATGCAC
CSN378	PtrpC-L3-fwd	AAGTGTAAU GCTAGTGGAGGTCAACACATCAATGC
CSN379	ampR-int-fwd	ACATGAUCCCCCATGTTGTGCAAAAAAG
CSN380	ampR-int-rv	ATCATGUAACTCGCCTTGATCGTTG
CSN381	argB-L3-alt fwd	AAGTGTAAU ACTAGGTAATATCGCGTGCATTCCG

CSN382	pyrG-L3-alt-fwd	AAGTGTAAU TCCTCGTGTACTGTGTAAGCGCCC
CSN385	Anid-Ptef-Pac-regen-fwd	AGGGTTUAATTAAGACCTCAGC CGAGACAGCAGAATCACCGC
CSN386	Ama-alt-out-pac-regen-rv	AAACCCUCAGC GGAAACAGCTATGACCATGAGATCTTCCCG
KBRP087	AMA1-int-rv	ATTGGGGUACTAACATAGCCATCAAATGCC
KBRP088	AMA1-int-fwd	ACCCCAAUGGAAACGGTGAGAGTCCAGTG
CSN351	Cas9-D10A-rv	ATGGCCAGU CCGATGCTATACTTCTTGTCCAT
CSN352	Cas9-D10A-fwd	ACTGGCCAU TGGAACGAACTCGGTTGGT
CSN353	Cas9-H840A-rv	ATGGCGU CAACATCGTAGTCCGACAACCG
CSN354	Cas9H840A-fwd	ACGCCAU CGTCCCGCAATCCTTCCTT
Vectors with alternative sgRNAs		
CSN521	gRNA-PS13-rv	AGCTTACUCGTTTCGTCCTCACGGACTCATCAGTCGGCGCGGTGATGTCTGCTCAAGCG
CSN522	gRNA-PS13-fwd	AGTAAGCUCGTC TCGGCGGAGTATCATAATCG GTTTTAGAGCTAGAAATAGCAAGTTAAA
CSN522	gRNA-PS14-fwd	AGTAAGCUCGTCTCGGCGGAGTATCATAATCGGTTTTAGAGCTAGAAATAGCAAGTTAAA
CSN536	gRNA-PS14-rv	AGCTTACUCGTTTCGTCCTCACGGACTCATCAGGGAGTACGGTGATGTCTGCTCAAGCG
CSN537	gRNA-PS14-fwd	AGTAAGCUCGTCGGAGTATCATAACATCGGTTTTAGAGCTAGAAATAGCAAGTTAAA

CSN574	gRNA-PS20-rv	AGCTTACUCGTTTCGTCCTCACGGACTCATCAG ATCATA CGGTGATGTCTGCTCAAGCG
CSN575	gRNA-PS20-fwd	AGTAAGCUCGTC ATCATAATCGAGGTTGAGTC GTTTTAGAGCTAGAAATAGCAAGTTAAA
***yA* repair templates**		
CSN518	yA-repair-90bp	TTGCTACGGACTGGACATATCTTACTTCGGCGGAGTATCATAACAT CGAGGTTGAGTCTGGCTATAACGTCTTGTAAGCTCGATTTTTCC
CSN519	yA-2bp-del-90bp	CTTGCTACGGACTGGACATATCTTACTTCGGCGGAGTATCATAATC GAGGTTGAGTCTGGCTATAACGTCTTGTAAGCTCGATTTTTCCT
CSN520	yA-repair-60-bp	GACATATCTTACTTCGGCGGAGTATCATAACATCGAGGTTGAGTCT GGCTATAACGTCTT
CSN568	yA-rep-500-fwd	GGGTTTAAU TGTCATCAATAACCTTCCCTTCAA
CSN569	yA-rep-500-rv	GGTCTTAAU AACTGTCCGTCAACTTCATATACCC
CSN570	yA-rep-1000-fwd	GGGTTTAAU GCAAATCCAGTAGGAAAATACGGT
CSN571	yA-rep1000-rv	GGTCTTAAU GTTTTCTGTCGTGAGGGCGTA
CSN572	yA-rep-2000-fwd	GGGTTTAAU GCACCATCGACCGTTTGATAT
CSN573	yA-rep-2000-rv	GGTCTTAAU GTCTCCGACGCAATCACCAT

Chapter 5: Conclusion and perspectives

This thesis presents much of the research done during my PhD, for which the overarching theme has been to explore different ways in which DNA repair can be manipulated, to enhance gene targeting.

In chapter **two**, I explored a method for making multiple transient gene disruptions, by the means of cre-lox recombination. This was achieved by inserting a dominant selection marker, flanked by loxP sites into an intron of a gene to be disrupted. While inserted the insert will prevent correct splicing and thus disable protein function, however once excised by cre-lox recombination only a single loxP scar is left and this is small enough to be tolerated in the intron. This was demonstrated to be a successful strategy for *yA* and *pyrG*. Unfortunately, whether it was likewise applicable for *nkuA* was not tested during this study. Nonetheless, there is reason to believe this will also work for *nkuA,* providing an alternative to transient inactivation by direct repeat recombination in which multiple genes can be inactivated simultaneously.

In chapter **three**, the focus changed from finding a smarter way of transiently disabling NHEJ combined with other genes to circumvent the problem of having to disable NHEJ, to have high frequencies of gene targeting in the first place. The genome editing technology CRISPR-Cas9, originating from bacteria and archaea, but recently adapted to a wide range of eukaryotes, showed promises that it could similarly be of great use in filamentous fungi. A vector set was created, based on the AMA1 element, allowing for transient expression in a wide range of *Aspergilli* and also some species outside the genus, such as some *Penicillium* and *Talaromyces* species. A codon optimized Cas9, optimized to *A. niger* and tagged with an C-terminal SV40 nuclear localization signal and combined with the AMA1 element and one of four selection markers, *pyrG*, *argB*, *hph* or *ble* makes the core of each vector. The main challenge was finding a way to express the other component of the CRISPR-Cas9 system, the sgRNA. By flanking the sgRNA with ribozymes, it was possible to use RNA polymerase II promoter for expression of the sgRNA [1], while normally an RNA polymerase III promoter is necessary. With a functional CRISPR-Cas9 system in place, I demonstrated that it could be used to both introduce small indels in six different NHEJ proficient species by faulty NHEJ repair, but also that it could be used to greatly enhance frequencies of repair by homologous recombination, similar to that on an NHEJ deficient strain, if a donor DNA template was co-transformed.

In chapter **four**, three themes were explored, building upon the CRISPR-Cas9 platform presented in chapter three. First four alternative promoters were tested for their ability to be used for sgRNA expression. Unfortunately, none of them worked in *A. nidulans*. Interestingly for three of them, very similar strategies have been successfully used in other filamentous fungi. One important thing to notice with the ribozyme based expression strategy presented in chapter three, since it is based on RNA polymerase II promoters, is that it also allows for the use of inducible promoters, while RNA polymerase III promoters are usually coding for household genes and as such are constitutively expressed. Next two strategies for limiting off-targeting effects were tested. One was based on truncating the protospacer, instead of using 20 bp, 17 bp was used, however this shortened protospacer resulted in abnormal sporeless colonies, strongly suggesting that this is not an appropriate strategy to use in *A. nidulans*. Also tested was using paired nickases, using two sgRNAs, causing two nicks in proximity, rather than causing a single double strand break. Results showed that while paired nickases are unsuitable to do mutagenesis, likely because the sticky ends are far more easily repaired in an error-free way than the blunt ends caused by wildtype Cas9. For repair by HR, however it showed promises. Two types of colonies resulted from transformations. A subpopulation showed a small and spore-less phenotype, and a more normal type of colony. While the number of "normal" colonies were low, the ratio between green and yellow colonies suggested it can be a viable strategy, if off-targeting ends up being a concern. The last area of investigation was whether CRISPR-Cas9 could be used to allow for insertions, deletions and point mutations without having to integrate aa accompanying selection marker. The main conclusions are that it is certainly possible, but that NHEJ deficiency greatly enhances correct integration, similar to how it enhances conventional gene targeting. Furthermore, I demonstrated that short single-stranded DNA oligos could be used as a repair templates, with varying amount of success. While there is still some optimizations to be done, the fact that various genetic modifications can be performed without the integration of a selection marker, is an important finding, since it is a prerequisite for multiplexing without being limited by the number of selection markers,.

The adaptation of the CRISPR-Cas9 genome editing tool for the use in Aspergilli and other filamentous fungi brings many exciting prospects, besides those outlined in this thesis. CRISPR-Cas9 is not the first genome editing technology, however what is separating it from its predecessors is the low barriers of entry. Meganucleases, while useful for some purposes are in general too specific to be useful for genome editing. Zinc-finger nucleases and to a lesser degree TALENs are expensive and time-consuming to develop, and for difficult to engineer organism, such as

mammalian cell systems, the effort to develop a ZFN or a TALEN to target a specific locus, might be worth it. Aspergilli however are significantly easier to engineer, while not as easy as some of the classical cell-factories such *S. cerevisiae and E. coli*, and so far the efforts have not really been worth it and thus explored, except for a recent paper describing the use of TALENs in *P.oryzae* [2]. However with CRISPR-Cas9, once a system is established, the cost per target is low enough, that even for organisms which are very genetically amendable, the technology has proven useful. Similarly, for Aspergilli it has the possibility to transform the conventional strategies for doing genetic manipulations.

Of course, CRISPR-Cas9 is extremely interesting from the perspective of genome editing, having the potential to change the paradigms of how we perform genetic engineering. However, another interesting application is for use in gene regulation. By inactivating of both cleavage domains, the resulting protein, dCas9 is a programmable DNA binding protein, which can be fused to other domains, e.g. the synthetic VP64 [3] acting as transcriptional activator, or repressing domains. The filamentous fungi, including Aspergilli represents a huge reservoir of interesting secondary metabolites, however many clusters are silent under laboratory conditions, and as such synthetic transcriptional activators could be a valuable tool. While the focus of this thesis have been on the genome editing aspects of CRISPR-Cas9, it lays a solid foundation to build tools for use in gene regulation rather than gene editing.

Finally, it is interesting to note that while the focus both in this thesis but also in general when it comes to this new generation of genome editing based on RNA-guided nucleases, has to a very high degree been on the *S. pyogenes* Cas9. Homologs of the *S. pyogenes* Cas9 have be identified, characterized, and utilized, but recently other proteins distinct from Cas9 have been investigated for their ability to be used for genome editing. One of such systems is Cpf1 nuclease from *Francisella novicida* [4], which produce sticky ends rather than the blunt ends produced by Cas9, which can be a useful complement tool to Cas9 based genome editing. However, considering how prevalent CRISPR systems for immune defense are in both bacteria and archaea, this is probably just the tip of the iceberg and more valuable additions to the genetic engineering toolbox could be out there.

Ultimately, this thesis presents some important advances in the application of genome editing technologies in *Aspergillus* and tools that can help facilitate faster and better genetic engineering strategies.

References

1. Gao Y, Zhao Y. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. J Integr Plant Biol. 2014;56: 343–349. doi:10.1111/jipb.12152

2. Arazoe T, Ogawa T, Miyoshi K, Yamato T, Ohsato S, Sakuma T, et al. Tailor-made TALEN system for highly efficient targeted gene replacement in the rice blast fungus. Biotechnol Bioeng. 2015;112: 1335–1342. doi:10.1002/bit.25559

3. Beerli RR, Segal DJ, Dreier B, Barbas CF. Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. Proc Natl Acad Sci U S A. 1998;95: 14628–33. doi:10.1073/pnas.95.25.14628

4. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, et al. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell. Elsevier Inc.; 2015;163: 759–771. doi:10.1016/j.cell.2015.09.038